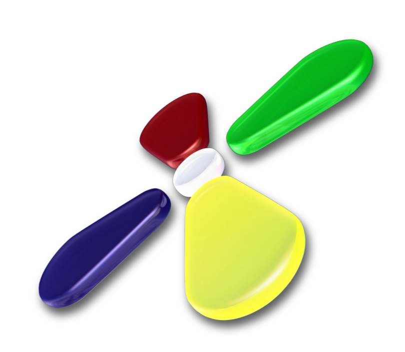
Das simon Handbuch



Das simon Handbuch

by Peter H. Grasch

simon ist eine Open-Source Spracherkennungslösung.

Inhaltsverzeichnis

List of Tables

Chapter 1. Einleitung

simon ist das Haupt-Frontend zu der simon Open-Source Spracherkennungslösung. Es ist ein simond Client und bietet ein grafisches Interface um das Sprachmodell und die Kommandos zu verwalten. Außerdem kann simon mit einer Vielzahl von Kommandos auf das Erkennungsergebnis reagieren, das er vom Server (simond) erhält.

Im Gegensatz zu existierenden kommerzieller Software bietet simon einen einzigartigen do-it-yourself Ansatz für Spracherkennung. Statt vordefinierten, bereits trainierten Sprachmodellen wird simon mit keinem Sprachmodell ausgeliefert. Stattdessen bietet es eine einfache Benutzeroberfläche für Endbenutzer um das Sprach- und Akustikmodell selbst zu erstellen.

Zusätzlich kann der Endbenutzer die Anwendungsfälle von anderen Benutzern einfach herunterladen und seine eigenen einfach mit anderen Benutzer austauschen.

Die aktuelle Version kann dazu benutzt werden, um Command & Control Lösungen zu verwirklichen, die besonders bei der körperlich eingeschräkte Zielgruppe Anwendungen finden dürfte. Eine freie Diktation ist aufgrund der großen Menge an dafür notwendigen Trainingsdaten in der derzeitigen Version nicht vorgesehen.

Wegen simons Architektur kann die selbe Version für alle Sprachen und Dialekte verwendet werden. Der Benutzer kann sogar verschiedene Sprachen innerhalb eines Modells mixen, sollte dies gewünscht sein.

Chapter 2. Übersicht

Architektur

Die Erkennungsarchitektur von simon besteht aus drei Anwendungen.

- simon

  Das grafische Interface.

  Es agiert als Client zum simond Server.

- simond

  Der Erkennungsserver.

- ksimond

  Das grafische Frontend für simond

Diese drei Komponenten bilden eine Server / Client Anwendung. Das bedeutet, dass ein Server (simond) mehrere Clients (simon; Diese Anwendung) bedient, welche über das Netzwerk angebunden sind. KSimond ist nur ein Frontend für simond was bedeutet, dass es keine zusätzliche Funktionalität zum System einbringt aber einen Weg bietet mit simond über eine grafische Benutzeroberfläche zu interagieren.

Neben simon, simond und ksimond sind andere, speziellere Applikationen auch Teil der integrierten simon Distribution.

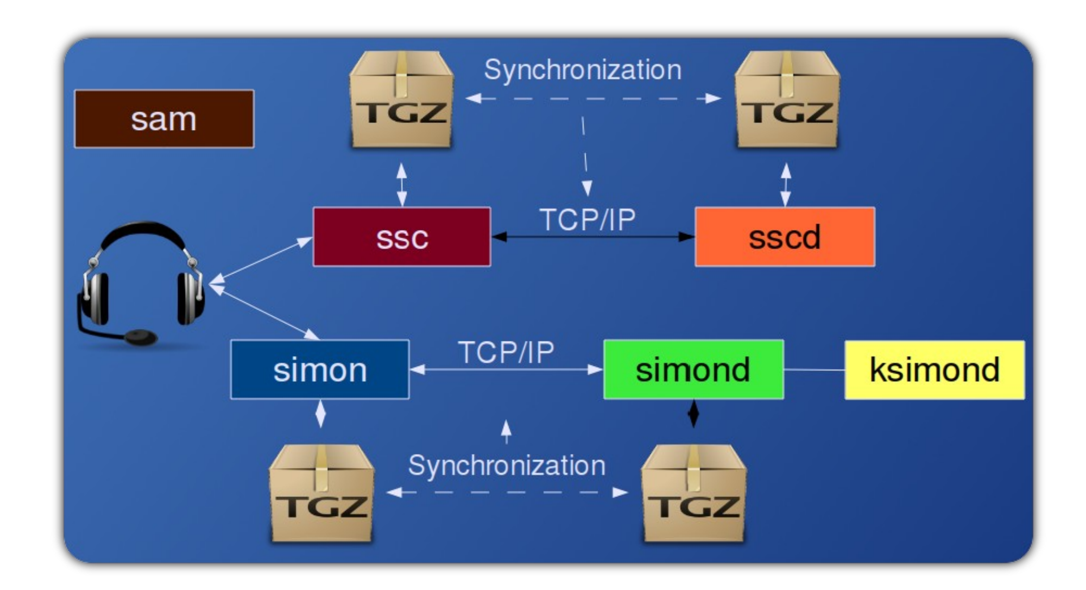
- sam

  Bietet direkten Zugang zum Sprachmodell um dieses zu testen oder zu generieren.

- ssc / sscd

  Diese beiden Applikationen können verwendet werden um eine große Menge an Sprachdaten von verschiedensten Personen leicht aufzunehmen.

Bitte sehen Sie die einzelnen Handbücher dieser Applikationen für mehr Details.



simon wird verwendet um eine RepräsentationIhrer Aussprache und Ihres Wortschatzes zu verwalten. Diese Repräsentationwird dann an den Server simond geschickt, welcher sie in ein benutzbares Sprachmodell übersetzt ("kompiliert").

simon nimmt mit Hilfe von einem Mikrofon dann Ihre Stimme auf und sendet diese Aufnahme an den Server welcher die Erkennung auf diesen Datenstrom ausführt. simond sendet dieses Erkennungsergebnis dann zurück an den Client (simon).

simon benutzt dieses Erkennungsergebnis dann um Kommandos auszuführen wie beispielsweise das Öffnen von Programmen oder URLs.

simond identifiziert die Verbindung mit einem Benutzernamen und einem Passwort die komplett unabhägig sind vom zugrunde liegenden Betriebssystem und dessen Benutzerkonten. Standardmäßig kommt simond mit einem Standardbenutzer in sowohl simon als auch simond so das der typische Anwendungsfall von einem simon Client auf einem simond Server ohne zusätzliche Konfiguration funktionieren wird.

Jeder simon Client meldet sich am Server mit einem Benutzernamen und einem Passwort an das einen eindeutigen Benutzer mit ein eindeutiges Sprachmodell identifiziert. Jeder Benutzer verwaltet sein eigenes Sprachmodell aber kann es von verschiedenen Computern (verschiedene physische simon Instanzen) verwenden einfach indem er den selben simond Server verwenden. Eine simond instanz kann natürlich mehrere Benutzer verwalten.

Wenn Sie einen Server zum Internet hin öffnen wollen oder mehrere Benutzer auf einem simond Server verwenden wollen, müssen Sie den Server entsprechend konfigurieren. Bitte sehen Sie das simond manual (help:/simond) für Details.

Spracherkennung: Hintergrundwissen

Spracherkennungssysteme verwandeln Spracheingabe (von einem Mikrofon) und versuchen diese in einen geschriebenen Text umzuwandeln. Um das zu erreichen, benutzen sie eine statistische Repräsentation der gesprochenen Sprache. Um es einfach auszudrücken: Der Computer lernt wie die Wörter - oder besser gesagt die Laute die die Wörter ausmachen - klingen.

Ein Sprachmodell besteht aus zwei Teilen (im Deutschen wird leider nicht zwischen "Sprache" im Sinne von "Sprechweise" und der "Sprache" im Sinne von Deutsch, Englisch, etc. unterschieden; Deshalb besteht das Sprachmodell ("speech model") aus einem Sprachmodell ("language model") und einem Akustikmodell ("acoustic model")):

- Sprachmodell
- Akustikmodell

Sprachmodell

Das Sprachmodell definiert das Vokabular und die Grammatik die verwendet wird.

Für mehr Information sehen Sie bitte die Vokabular-Sektion und die Grammatik-Sektion.

Akustikmodell

Das Akustikmodell repräsentiert die Aussprache in einem maschinenlesbaren Format.

Betrachten wir das folgende Beispiels-Vokabular:

Table 2-1. Beispiels-Vokabular

Wort	Terminal	Aussprache
Computer	Nomen	k O m p j u: t @ r
Internet	Nomen	I n t @ r n E t
Mail	Nomen	m E I l
schließen	Verb	S l i: s @ n

Die Aussprache jedes Wortes besteht aus einzelnen Lauten die mit Leerzeichen getrennt sind. Das Wort "schließen" besteht bspw. aus den folgenden lauten:

- S
- l
- i:
- s
- @
- n

Das Akustikmodell nutzt die Tatsache, dass gesprochene Wörter genauso aus Lauten bestehen wie geschriebene Wörter aus Buchstaben bestehen. Darauf aufbauend können wir Wörter in einzelne Laute zerlegen und sie bei der Erkennung wieder "zusammenbauen". Diese Bausteine nennt man "Phoneme".

Das Akustikmodell beinhaltet Daten wie die einzelnen Phoneme und nicht ganze Wörter vom Anwender ausgesprochen werden. Das Trainingsmaterial eines Wortes beeinflusst deshalb auch die Erkennungsrate ähnlicher Worte positiv.

Das bedeutet, dass wenn zum obigen Beispielvokabular nun noch das Wort "Mailing" hinzugefügt wird, weiß das Akustikmodell bereits ungefähr wie es klingen wird weil das Wort aus den Phonemen "m", "E", "I", "l", "I", "N" besteht und wir die Phonemkombination "m", "E", "I", "l" ja bereits für "Mail" trainiert haben. Wird jetzt eines der beiden Wörter trainiert, verbessert sich die Erkennung beider

Wörter. Natürlich müssen aber auch die Phoneme "I" und "N" trainiert werden weshalb ein alleiniges Training vom Wort "Mail" nicht ausreicht um "Mailing" zu erkennen.

Um das Akustikmodell zu trainieren (in anderen Worten, dem Computer zu erklären wie die Phoneme von Ihnen ausgesprochen werden) müssen Wörter aus Ihrem Sprachmodell "trainiert" werden. Das bedeutet, dass simon das Wort anzeigt und Sie dieses laut vorlesen. Weil simon das Wort aus seinem Vokabular bereits kennt, weiß er aus welchen Phonemen es besteht. Damit kann er dann Ihre persönliche Aussprache dieser Phoneme "erlernen".

Szenarios

Ein Szenario bestimmt einen kompletten Anwendungsfall von simon. Um Firefox zu steuern, zum beispiel, kann man einfach das Firefox Szenario installieren.

Jedes Szenario beinhaltet die folgenden Komponenten:

- Vokabular
- Grammatik
- Trainingstexte
- Kommandos

Szenarios decken nur das Sprachmodell (language model) des Erkennungssystem ab, das Akustikmodell (acoustic model) ist dabei vollkommen unabhägig.

In den meisten Fällen sind Szenarios so entwickelt um optimal mit einem gewissen Basismodell zusammenzuarbeiten um Probleme mit dem Phonemset zu vermeiden.

Weil Szenarios nicht direkt an ein Akustikmodell gebunden sind können Sie weitergegeben werden und ohne Probleme mit anderen simon Anwendern ausgetauscht werden . Um diesen Anwendungsfall möglichst gut abzubilden, wurde eine Kategorie für simon Szenarios auf kde-files.org (http://kde-files.org/index.php?xcontentmode=692) eingerichtet. Dort können Szenarios, die intern simple Textdateien im XML Format sind, komfortabel hoch- und heruntergeladen werden.

Für Informationen wie Sie Szenarios in simon verwenden können, sehen Sie bitte die Szenario Sektion im simon Verwenden Kapitel.

Basismodelle

Basismodelle sind bereits generierte, meist sprecherunabhägige Akustikmodelle die mit simon verwendet werden können.

Mit Basismodellen kann ein Anwender die benötigte Zeit für ein individuelles Training stark reduzieren. Bei Verwendung von statischen Basismodellen (siehe unten) ist eine Installation des HTKs nicht notwendig.

Basismodelle die von simon verwendet werden können bestehen aus vier Dateien die Sie in einem Archiv finden werden wenn Sie die Modelle von der jeweiligen Webseite herunterladen.

- hmmdefs
- tiedlist
- macros

- stats

Die letzten beiden Dateien, macros und stats, sind nicht notwendig wenn Sie statische Basismodelle verwenden und können in diesem Fall mit leeren Dateien ersetzt werden sollten Sie nicht verfügbar sein.

Wo bekommen Sie Basismodelle

Um diese Liste aktuell zu halten, sehen Sie bitte die Liste in unserem Wiki (http://www.simon-listens.org/wiki/index.php/English:_Base_models#Where_to_get_base_models).

Verschiedene Typen von Basismodellen

Es existieren drei typen von Basismodellen:

- Statisches Basismodell
- Adaptiertes Basismodell
- Benutzergeneriertes Modell

Für Informationen wie Sie Basismodelle mit simon verwenden können, sehen Sie bitte die Basismodelle verwenden Sektion im simon Verwenden Kapitel

Statisches Basismodell

Statische Modelle verwenden einfach ein zuvor kompiliertes Akustikmodell ohne es zu bearbeiten.

Trainingsdaten die mit simon aufgenommen werden werden nicht verwendet um die Erkennungsrate zu erhöhen.

Dieser Typ von Modell *benötigt kein* HTK.

Adaptiertes Basismodell

Durch Adaption von bereits kompilierten Akustikmodellen können Sie die Erkennungsrate steigern indem Sie das Modell an Ihre Stimme anpassen.

Gesammelte Trainingsdaten werden in eine Adaptionsmatrix kompiliert die dann auf das ausgewählte Basismodell angewendet wird.

Dieser Typ von Modell *benötigt* ein installiertes HTK.

Benutergeneriertes Modell

Wenn Sie benutzergenerierte Modelle verwenden ist der Benutzer dafür verantwortlich, dass das Modell trainiert ist. Es wird kein Basismodell verwendet.

Die Trainingsdaten werden verwendet um das eigene Akustikmodell zu kompilieren. Dies erlaubt es ein System zu entwickeln das direkt Ihre eigene Sprache reflektiert.

Dieser Typ von Modell *benötigt* ein installiertes HTK.

Phonemset Probleme

Da der statistische Vergleich auf Phonemebene geschieht, beschreibt das Basismodell wie diese Phoneme klingen.

Ihr Szenario (Sprachmodell / Language model) hingegen beschreibt wie die Phoneme klingen aus denen die Wörter aufgebaut sind.

Damit diese Verbindung funktioniert, muss sowohl Ihr Szenario als auch Ihr Basismodell das selbe Phonemset verwenden.

Wenn Sie ein neues Szenario entwerfen ist es also eine gute Idee jenes Wörterbuch als Schattenvokabular zu verwenden das verwendet wurde um das Basismodell zu generieren. Auf diese Weise wird simon automatisch die "richtigen" Phoneme vorschlagen wenn Wörter hinzugefügt werden.

Wenn Sie Szenarios verwenden die für ein anderes Phonemset entwickelt wurden (anderes Basismodell) werden Sie unter umstäden beim Starten der Erkennung einen Fehler erhalten der betroffene Phoneme und damit assoziierte Wörter anzeigt. Um dies zu reparieren, transkripieren Sie ihre Wörter entsprechend des verwendeten Phonemset des Basismodells oder verwenden Sie ein anderes Basismodell oder benutzen Sie ein benutzergeneriertes Modell.

Chapter 3. Richtlinien

Dieses Kapitel listet ein paar grundlegende Richtlinien die in unterschiedlichen Teilen von simon relevant sind.

Aufnahmen

Wenn Sie benutzergenerierte oder adaptierte Modelle verwenden wird simon basierend auf transkripierten Sprachaufnahmen von der Stimme des Benutzers auf sein Akustikmodell aufbauen. Deswegen ist die Qualität der aufgenommenen Trainingsdaten entscheidend für die Erkennungsrate.

Lautstärke

Es ist wichtig, dass Sie vor der Aufnahme die Lautstärke Ihres Mikrofons kontrollieren.

simon Kalibrierung

Die aktuelle Version von simon inkludiert einen einfachen Weg um sicherzustellen, dass die Lautstärke korrekt konfiguriert ist.



Standardmäßig wird die Lautstärkekalibrierung angezeigt bevor eine Aufnahme in simon startet.

Um die Lautstärke zu kalibrieren, sprechen Sie einfach den angezeigten Text nach.

Um die Kalibrierung zu überwachen, sehen Sie bitte die angezeigte Lautstärke und die Notiz darunter die sagt, ob Sie die Lautstärke heben oder senken sollten. Um die Lautstärke zu verädern, verwenden Sie bitte die Lautstärkekonfiguration Ihres Betriebssystems. Sobald Sie dort die Lautstärke verädert haben drücken Sie bitte in simon auf "Lautstärke geädert" um die Kallibrierung darüber zu informieren.

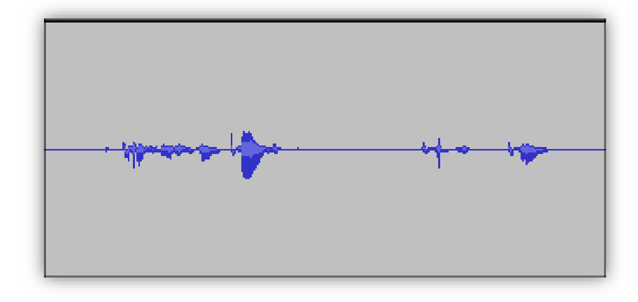
Versuchen Sie während der Kalibrierung so normal wie möglich zu sprechen. Schreien Sie nicht aber seien Sie auch nicht unnötig leise. Bitte bedenken Sie, dass Sie die besten Ergebnisse erzielen wenn alle Aufnahmen ungefähr gleich laut sind und Sie auch während der Erkennung gleich laut sprechen. Möglicherweise könnten Sie zu anderen Tageszeiten oder einer anderen Stimmung lauter sprechen also sprechen Sie in diesem Test lieber etwas lauter um diese Möglichkeit einzukalkulieren. Grundsätzlich sind zu leise Aufnahmen viel besser als zu laute, da es dann zu Verzerrungen kommen kann.

In den Einstellungen von simon lässt sich sowohl der angezeigte Text als auch die Lautstärken die als korrekt bzw. inkorrekt markiert werden konfigurieren. Wird der dortige Text leer gelassen (Standard) wird ein Standardtext angezeigt. In der Konfiguration können Sie die Lautstärkenkalibrierung auch komplett deaktivieren. Sehen Sie Training Sektion für mehr Details.

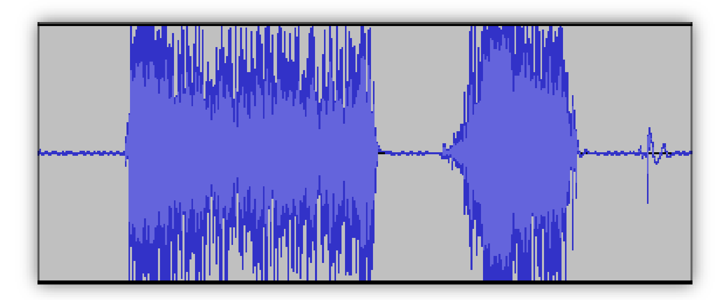
Audacity Kalibrierung

Ein alternativer Weg um sicherzustellen, dass die Lautstärke richtig eingestellt ist, ist ein Audio Tool wie das freie Audacity (http://audacity.sf.net) zu benutzen.

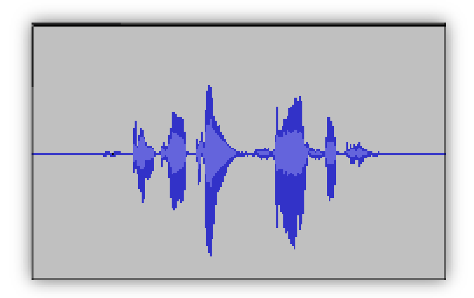
Zu leise:



Zu laut:



Optimale Lautstärke:



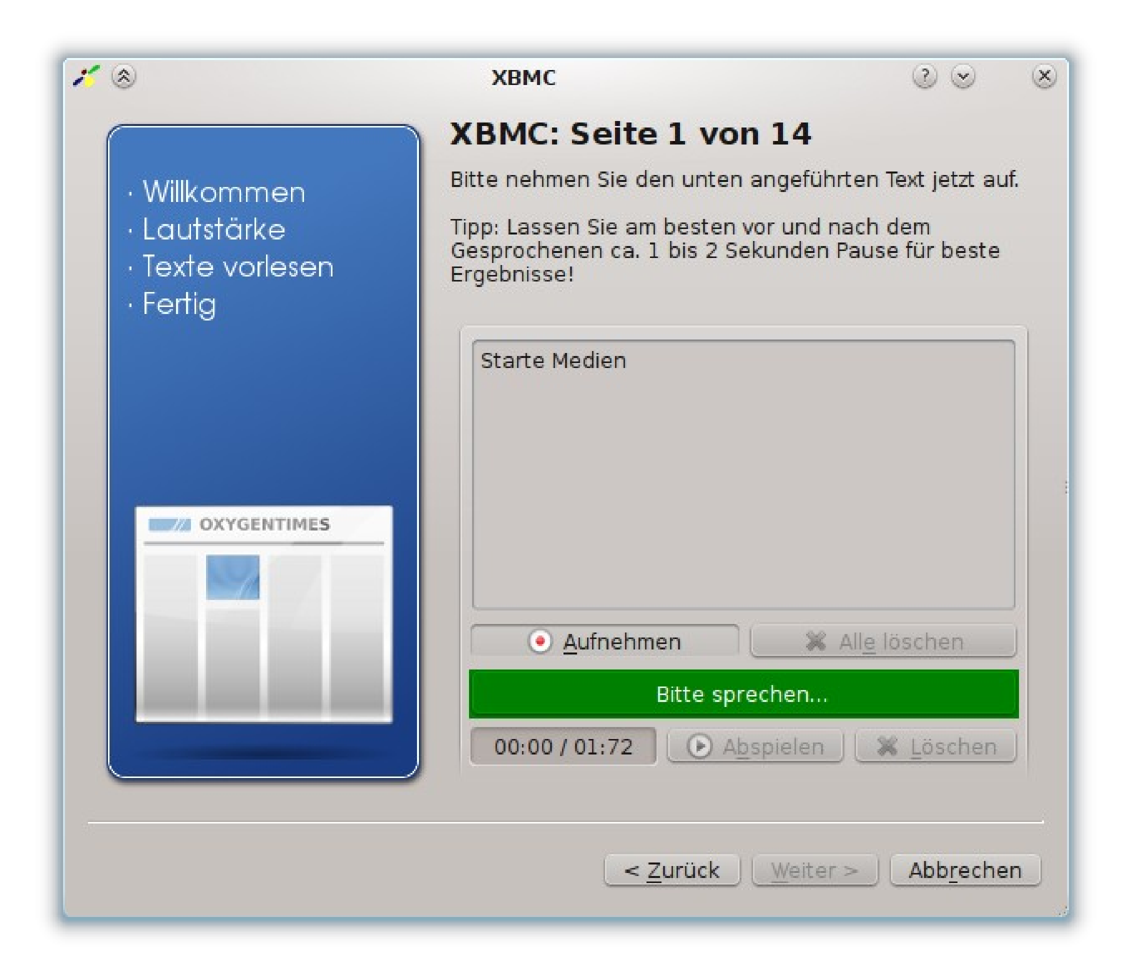
Stille

Um simon mit der automatischen Segmentierung zwischen den Wörtern zu helfen, empfehlen wir vor und nach dem Sprechen ca. ein bis zwei Sekunden "Stille" aufzunehmen.

Die aktuelle simon Version inkludiert eine grafische Warteanzeige während dem Training. Diese Nachricht wird den Benutzer bitte eine Sekunde zu warten:



... bevor er gebeten wird zu sprechen:



Diese Methode des visuellen Feedbacks hat sich als besonders wertvoll erwiesen wenn mit Personen aufgenommen wird die den angezeigten text nicht selbst lesen können und deshalb auf Moderatoren angewiesen sind die Ihnen sagen was Sie sprechen sollen. Die farbige Anzeige hilft Ihnen zu erkennen, wann Sie sprechen sollen statt direkt nach dem Moderator zu sprechen oder auf umstädliche Handsignale angewiesen zu sein.

Mikrofon

Damit simon gut funktionieren kann ist ein hochqualitatives Mikrofon empfohlen.

Aber sogar mit relativ günstigen Headsets (um 30 â¬) erreicht sehr gute Ergebnisse - viel bessere als ein möglicherweise integriertes Mikrofon.

Für maximale Kompatibilität empfehlen wir USB Headset die in der Regel die notwendige Samplerate von 16kHz direkt unterstützen. Sie sind außerdem gut von sowohl Microsoft Windows als auch von GNU/Linux unterstützt und benötigen keine speziellen, properitären Treiber.

Aufnahmen Qualitätskontrolle

simon wird jede Aufnahme mit einigen Kriterien überprüfen um sicherzustellen, dass die Aufnahmen nicht fehlerhaft oder von schlechter Qualität sind.

Wenn simon eine problematische Aufnahme erkennt, wird er den Benutzer warnen und ihn auffordern die Aufnahme zu wiederholen.

Derzeit überprüft simon die folgenden Kriterien:

- Höchste Lautstärkepegel

  Wenn die Lautstärke zu hoch ist und das Mikrofon beginnt zu "übersteuern" (Übersteuern auf Wikipedia (http://de.wikipedia.org/wiki/%C3%9Cbersteuern_%28Signalverarbeitung%29)). simon wird automatisch eine Warnung anzeigen die den Benutzer auffordert die Lautstärke des Mikrofons zu senken.

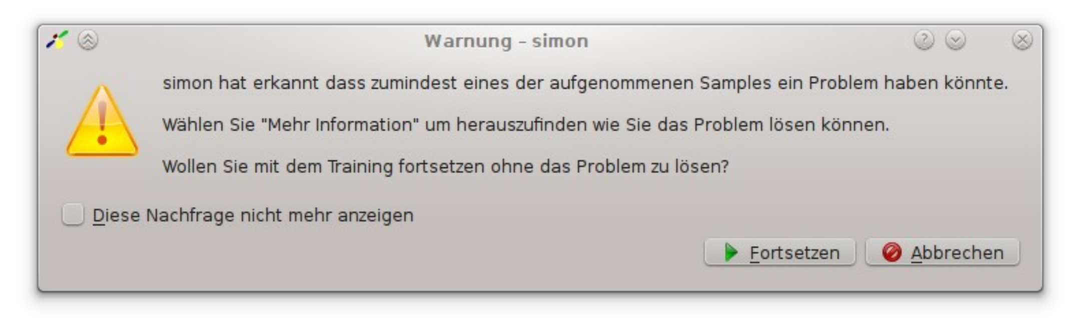
- Signal zu Rauschen Verhältnis (engl.: Signal to noise ratio; SNR)

  simon wird automatisch das Signal / Rausch Verhältnis für jede Aufnahme bestimmen. Ist dieses Verhältnis unter einem konfigurierbaren Level, wird eine Warnung angezeigt.

  Der Defaultwert von 5000 % bedeutet, dass simon nur Samples als "korrekt" ansieht, deren Spitzenpegel 500 mal lauter ist als der niedrigste Lautstärke (niedrigste Lautstärke wird als niedrigster Durchschnitt von jeweils 50 Millisekunden errechnet).

  Oft ist dies ein Problem von sehr schlechten Mikrofonen, hohen Leveln von Umgebungsgeräuschen oder eine kleine Mikrofonlautstärke in Verbindung mit einem "Signalboost" im Systemmixer.

SNR Warnungen werden naturgemäß auch von leeren Aufnahmen ausgelöst. Dieser Dialog wird bspw. Angezeigt wenn auf den Mehr Information Knopf gedrückt wird (sichtbar im Hintergrund).



Chapter 4. simon benutzen

Die folgende Sektion wird die "normale" Benutzung von simon beschreiben.

Das simon Hauptfenster



Das simon Hauptfenster bietet schnellen Zugriff auf die wichtigsten Funktionen durch die Haupt-Werkzeugleiste.

Es existieren 9 Aktionen in der Toolbar:

- simond Verbindung

  Dieses Menüobjekt hat verschiedene Zustäde:

  - Verbinden

    Wenn simon nicht zum simond verbunden ist, ist die Option mit "Verbinden" beschriftet. Wenn die Option aktiviert wird, wird simon eine Verbindung zum simond aufbauen und die Option wechselt in den "Verbinde"-Status.

    Wenn dies das erste Mal ist, das eine Verbindung zum simond Server aufgebaut wird, kann es sein, dass simon Sie nach einem Benutzername und einem Passwort fragt. Wenn Sie noch keinen simond

Benutzer angelegt haben, sollten Sie das jetzt machen (sehen Sie das simond Handbuch (help:/simond) für Details) bevor Sie fortfahren. Haben Sie bereits einen Benutzer in simond angelegt, geben Sie seine Daten bitte in das Anmeldefenster ein. Wenn Sie auswählen, das Passwort zu speichern wird Sie simon beim nächsten Verbindungsaufbau nicht mehr nach den Anmeldedaten fragen. Sie können sie aber natürlich jederzeit in der Server Konfiguration ädern.

  - Verbinde

    Wenn simon gerade eine Verbindung zum simond aufbaut, ädert sich der Text in "Verbinde" und der Knopf ist heruntergedrückt. Wenn die Option dann ausgewählt wird, wird der Verbindungsaufbau abgebrochen und die Option wechselt zurück in den "Verbinden"-Status.

  - Aktivieren

    Wenn simon eine Verbindung mit einem simond hergestellt hat, wird sich der Text in "Aktivieren" ädern und wird nicht heruntergedrückt sein. Wenn die Option aktiviert wird (was automatisch passiert wenn die Erkennung verfügbar ist und die Option "Erkennung automatisch starten sobald verfügbar" aktiviert ist) wird simon versuchen die Erkennung zu starten.

    Eine Option, die Verbindung zum simond zu trennen ("Verbindung trennen") ist verfügbar über das kleine Häkchen neben der Option.

  - Aktiviert

    Wenn simon eine Verbindung zum Server aufgebaut hat und die Erkennung starten konnte, wird die Option mit "Aktiviert" beschriftet und heruntergedrückt sein. Wenn die Option jetzt aktiviert wird, wird simon die Erkennung stoppen und somit in den "Aktivieren"-Status zurückfallen. Die Verbindung zum Server wird dadurch aber nicht getrennt.

    Eine Option, die Verbindung zum simond zu trennen ("Verbindung trennen") ist verfügbar über das kleine Häkchen neben der Option.

- Wort hinzufügen

  Zeigt den Wort Hinzufügen Assistenten an.

- Vokabular

  Zeigt das Vokabular an.

- Grammatik

  Zeigt die Grammatik an.

- Training

  Zeigt das Training an.

- Kommandos

  Zeigt die Komamndos an.

- Synchronisieren

  Wenn simon zum simond verbunden ist, ist diese Option verfügbar.

  simon erstellt die Eingabedateien für das Sprachmodell welche dann kompiliert und vom simond Server für die Erkennung verwendet werden (siehe auch das Kapitel Architektur für mehr Details).

  Der Prozess des sendens der Ausgabedaten, kompilieren dieser und erhalten des fertigen Sprachmodells wird "Synchronisation" genannt. Per Default wird simon eine Synchronisation

automatisch einleiten wenn er sich zum Server verbindet und jedes Mal, wenn sich die Eingabedateien für das Sprachmodell lokal ädern. Bitte sehen Sie die Sektion Synchronisation konfigurieren wenn Sie diese Einstellung ädern wollen.

Mit dieser Menüoption können Sie die Synchronisation jederzeit manuell auslösen.

- Szenario Auswahl

  Dieses Auswahlfeld erlaubt Ihnen das derzeit angezeigte Szenario anzupassen. Jede Untersektion (Vokabular, Grammatik, Kommandos, Training) wird dann das ausgewählte Szenario anzeigen. Das hier ausgewählte Szenario beeinflusst die Erkennung nicht.

- Szenarios verwalten

  Anzeigen des Szenarios verwalten Dialog. Hier können Sie Szenarios verwalten und die verfügbaren Optionen der Szenario Auswahlbox anpassen.

Das simon Hauptfenster kann jederzeit versteckt werden indem man auf das simon Logo im Systemabschnitt der Taskleiste (normalerweise der Bereich neben der Uhr) klickt. Wird das Symbol erneut gedrückt, wird simon wieder angezeigt.

Benötigte Resourcen für eine Funktionierende simon Konfiguration

> **Note:** Für mehr Information über Sprachmodelle sehen Sie bitte die Sektion Spracherkennung: Hintergrund im Übersichtskapitel.

Damit die Spracherkennung funktioniert, benötigt simon ein Sprachmodell (engl.: speech model).

Sprachmodelle (speech models) beschreiben wie Ihre Stimme klingt wie welche Wörter klingen, welche Wörter existieren und welche Kombinationen ("Sätze" oder "Strukturen") existieren.

Ein Sprachmodell besteht aus zwei Teilen:

- Sprachmodell (language model): Beschreibt welche Wörter existieren und welche Sätze grammatikalisch korrekt sind.
- Akustikmodell (acoustic model): Beschreibt wie welche Töne (Phoneme) klingen.

Sie benötigen beide Komponenten um die Erkennung in simon zu aktivieren.

Sprachmodell

In den meisten Fällen werden Sie nur die gewünschten Szenarios für Ihren Anwendungsfall einrichten um ihr Sprachmodell (language model) zu erstellen.

Um ihr eigenes Sprachmodell zu bauen, können Sie simon verwenden um Wörter hinzuzufügen / zu editieren / zu löschen und ihre Grammatikstrukturen zu verwalten.

Um das Hinzufügen von Wörtern leichter zu machen, können Sie ein Schattenlexikon importieren.

Akustikmodell

Um ein eigenes Akustikmodell zu erzeugen, können Sie einfach die Trainingstexte ein paar mal vorlesen die mit den Szenarios ausgeliefert werden die Sie installieren.

Wenn Sie ihre eigenen Szenarios erstellen können Sie leicht Trainingstexte selbst erstellen.

Sie können außerdem statische oder adaptierte Sprachmodelle verwenden um die Abhägigkeit des HTKs zu vermeiden oder die Erkennungsrate zu steigern.

Einrichtungsassistent

Beim ersten Start von simon wird eine Anleitung angezeigt die Sie durch die ersten Schritte mit simon führt.



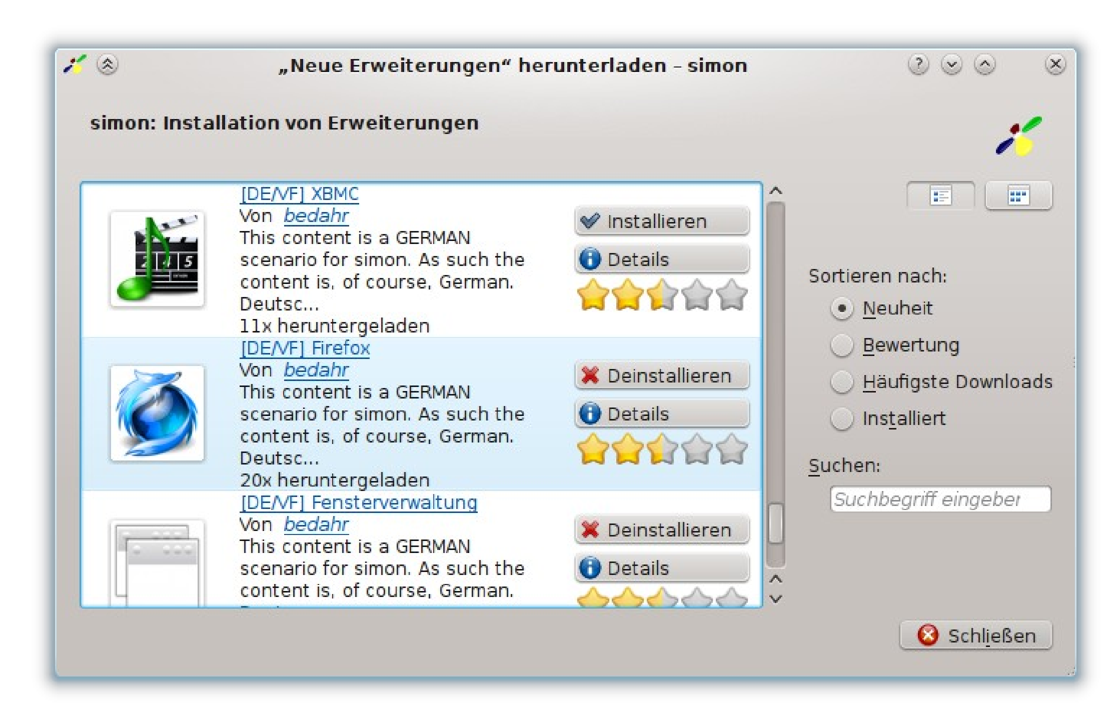
Die Konfiguration besteht aus fünf einfachen Schritten die hier aufgezeigt sind. Sie können jeden Schritt oder sogar den ganzen Assistenten überspringen. In diesem Fall wird das System mit Standardwerten eingerichtet.

Szenarios

In diesem Schritt können Sie Szenarios aus dem Internet herunterladen und direkt in simon importieren.



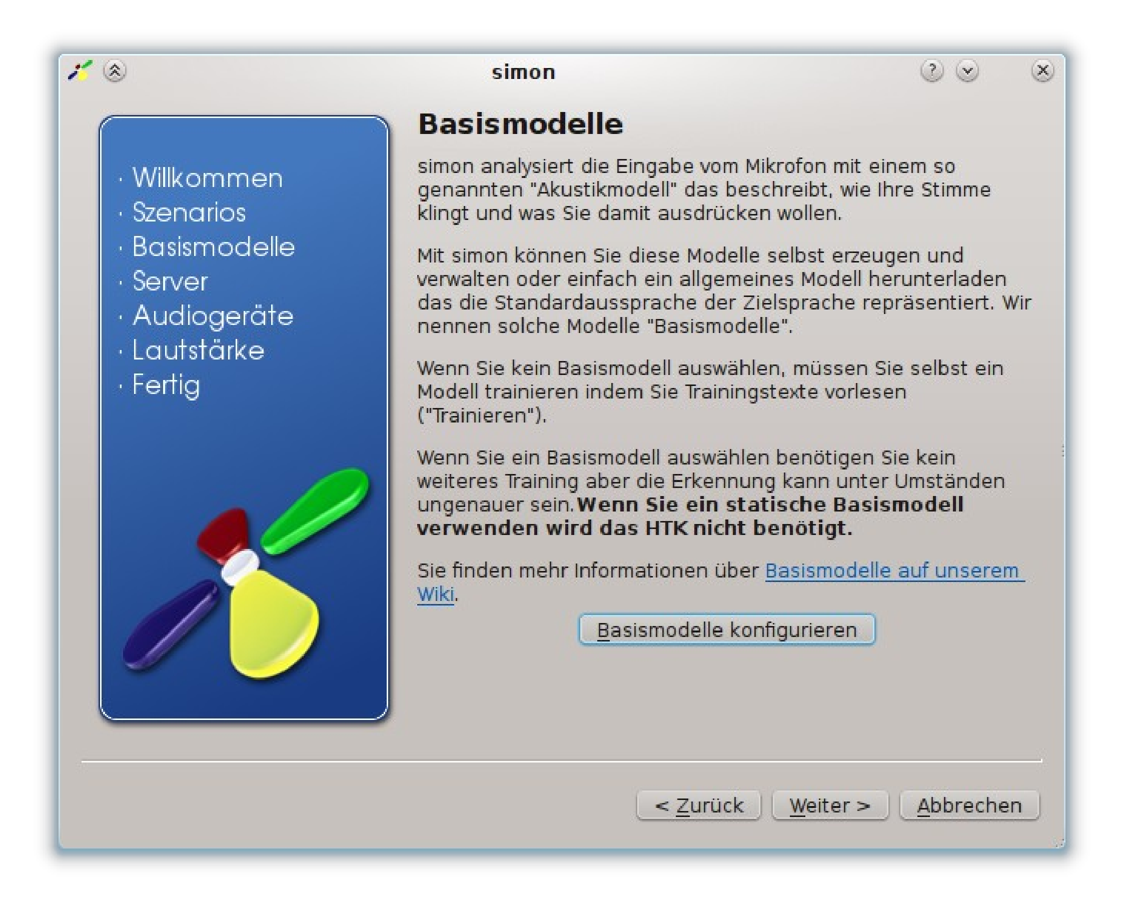
Wenn Sie Szenarios holen drücken, wird der Szenario Downloaddialog angezeigt.



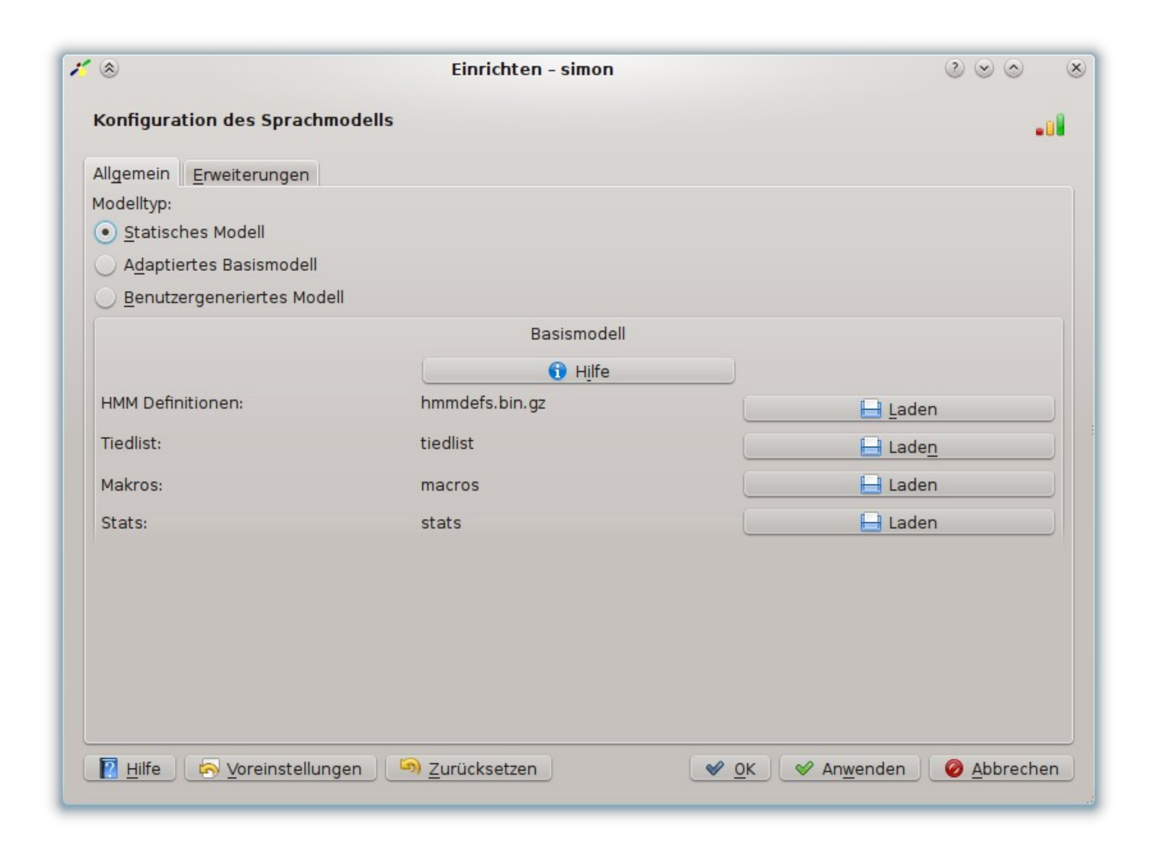
Wenn Sie hier Szenarios importieren (oder später über den Szenarioverwaltungsdialog) müssen Sie das Vokabular, die Grammatik und die Kommandos nicht selbst konfigurieren. Speziell für neue Anwender ist es deshalb empfohlen ein paar Szenarios hier herunterzuladen um zu sehen wie das System funktioniert.

Basismodelle

In diesem Schritt können Sie simon konfigurieren Basismodelle zu verwenden.



Die Konfigurationsseite die hier geöffnet wird ist die selbe die in der Verwendung von Basismodelle Sektion erklärt wird.



Nachdem der Einrichtungsassistent abgeschlossen ist oder abgebrochen wurde können Sie die Konfigurationsoptionen dieses Schritts in der simon Konfiguration jederzeit ädern.

Server

Intern ist simon eine Server / Client Anwendung. Wenn Sie Nutzen daraus durch eine netzwerkbasierten Verbindung ziehen wollen können Sie die Serveraddresse hier angeben.



Die Standardkonfiguration ist ausreichend für eine "normale" Installation und nimmt an, dass Sie einen lokalen simond Server verwenden der von simon automatisch gestartet und gestoppt wird.

Nachdem der Einrichtungsassistent abgeschlossen ist oder abgebrochen wurde können Sie die Konfigurationsoptionen dieses Schritts in der Server Konfiguration jederzeit ädern.

Aufnahmekonfiguration

Da simon Sound von einem oder meheren Mikrofonen aufnimmt müssen Sie simon zuerst sagen welche Geräte er für das Training bzw. die Erkennung verwenden soll.



simon kann ein oder mehrere Ein- und Ausgabegeräte für verschiedene Aufgaben verwenden. Sie können mehr informationen über simons Mehrgerätefähigkeiten in der Soundkonfiguration Sektion nachlesen.

Wenn Sie nicht zumindest ein Eingabegerät für die Erkennung definieren werden Sie simon nicht aktivieren können.

Wenn die Option Powertraining per Default aktiviert ist wird simon beim Trainieren automatisch die Aufnahme starten bzw. stoppen wenn der dazugehörige Prompt angezeigt bzw. versteckt wird. Diese option setzt nur den Standardwert der Benutzer kann diesen bei jeder Trainingssitzung wieder überschreiben.

Nachdem der Einrichtungsassistent abgeschlossen ist oder abgebrochen wurde können Sie die Konfigurationsoptionen dieses Schritts in der Soundkonfiguration jederzeit ädern.

Lautstärkeeinpegelung

Damit simon korrekt funktioniert muss die Mikrofonlautstärke korrekt eingestellt sein.



Für mehr Information sehen Sie bitte die Lautstärkeeinpegelung-Sektion und dasRichtlinien-Kapitel.

Szenarios

Diese Sektion beschreibt wie Sie Szenarios innerhalb von simon verwenden. Für generelle Informationen über Szenarios sehen Sie bitte das Hintergrund Kapitel.

Szenarios verwenden

Beginnend mit simon 0.3 wird jedes Wort das hinzugefügt wird zum aktuell aktiven Szenario hinzugefügt. Das selbe gilt für Grammatikstrukturen, Kommandos, etc.

Szenarios zu verwenden ist dann gleichzusetzen mit dem normalen verwenden von simon wie Sie es vielleicht schon in 0.2 gemacht haben.

Standardmäßig kommt simon mit einem leeren Szenario namens "Standard". Wenn Sie keine weitere Einstellungen vornehmen werden Ihre persönlichen Einstellungen dort gespeichert.

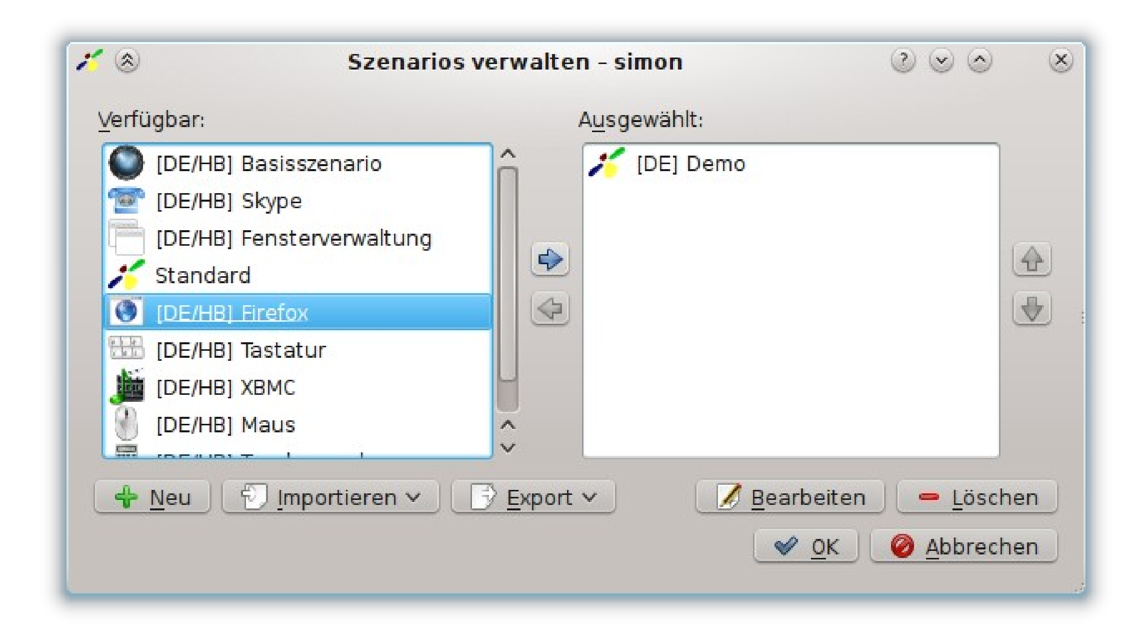
Um ein anderes Szenario zur Bearbeitung auszuwählen (z.B. vor dem Hinzufügen neuer Wörter), wählen Sie das Szenario einfach in der in der Auswahlliste im oberen rechten Eck des Hauptfensters aus.



Um die verfügbaren Optionen zu ädern, klicken Sie Szenarios verwalten oder verwenden Sie den Menüeintrag Szenarios > Szenarios verwalten.

Szenarios verwalten

Der Szenario verwaltungsdialog erlaubt Ihnen, Szenarios zu laden, neue zu importieren oder vorhandene in Dateien oder direkt in eine Onlinesammlung zu exportieren.

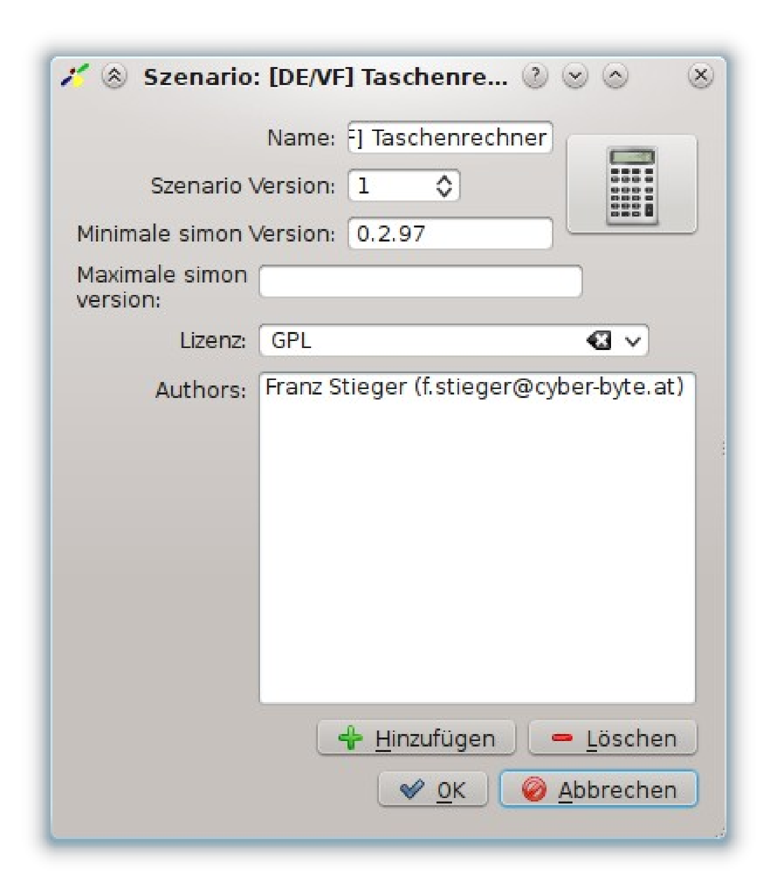


Um ein Szenario zu laden oder zu entladen, verwenden Sie die Knöpfe mit Pfeilsymbolen oder doppelklicken Sie das entsprechende Szenario in den Listen.

Mehr Informationen über die Szenarios selbst können Sie in dem Tip lesen der erscheint, wenn die Maus einige Zeit über ein Element halten.

Neues Szenario hinzufügen

Um ein neues Szenario hinzuzufügen, wählen Sie den Hinzufügen Knopf aus. Ein Dialog wird angezeigt.



Wenn Sie ein neues Szenario erzeugen, geben Sie Ihm bitte einen deskreptiven Namen. Um es später auf KDE files (http://kde-files.org/index.php?xcontentmode=692) hochzuladen würden wir Sie bitten das folgende Namensschema zu übernehmen: "[<Sprache>/<Basismodell>] <Name>". Wenn Sie zum Beispiel ein deutsches Szenario erzeugen das mit dem Voxforge Basismodell funktioniert um Mozilla Firefox zu kontrollieren wird das: "[DE/VF] Firefox". Ist Ihr Szenario nicht speziell auf ein Phonemset (Basismodell) zugeschnitten, können Sie den zweiten Tag entfernen: "[DE] Firefox".

Die Szenario Version ist nur eine inkrementelle Nummer die es leichter macht zwischen verschiedenen Versionen des selben Seznarios zu unterscheiden.

Wenn Ihr Szenario eine spezielle Funktion von simon verwendet (zum Beispiel ein neues Plugin), können Sie hier die minimal benötigte simon Versionsnummer angeben. Wurde die benötigte Funktionialität entfernt, lässt sich auch eine maximale Versionsnummer definieren.

Die Lizenz lässt sich über das Drop-Down Menü auswählen. Sie können auch beliebigen Text in das Eingabefeld eintragen.

Sie können Ihren Namen (oder Alias) zu der Liste der Szenarioautoren hinzufügen. Sie werden auch nach einer Kontaktmöglichkeit gefragt werden. Diese Kontaktmöglichkeit kann dazu verwendet werden den Autor über Änderungen in simon zu kontaktieren die das Szenariosystem betreffen, Fanmail, etc.

Wenn Sie keine Kontaktinformation hinterlegen wollen, geben Sie einfach einen Strich "-" ein.

Szenario editieren

Um Szenarios zu editieren, wählen Sie "Bearbeiten" vom "Szenarios verwalten" Dialog.

Der Dialog funktioniert exakt gleich wie der Szenario hinzufügen Dialog.

Szenario entfernen

Um ein Szenario zu löschen, wählen Sie das Szenario aus der Liste und klicken Sie den "Löschen" Knopf.

Weil Szenarios mit dem Erkennungsdienst synchronisiert werden, können Sie zuvor gelöschte Szenarios über das Modell Synchronisationsbackup wiederherstellen.

Szenario importieren

Szenarios können von einer lokalen Textdatei (simons XML Szenario Dateiformat) importiert werden oder direkt aus dem Internet heruntergeladen werden.

Wenn Sie Szenarios herunterladen, wird eine Liste der verfügbaren Szenarios von der simon Scenarios (http://kde-files.org/index.php?xcontentmode=692) Sektion der OpenDesktop Seite KDE-files.org (http://kde-files.org) heruntergeladen.



Wenn Sie ein Szenario erstellen das für andere simon Anwender von interesse sein könnte, laden Sie es bitte hoch damit es auch anderen simon Anwendern zugute kommen kann.

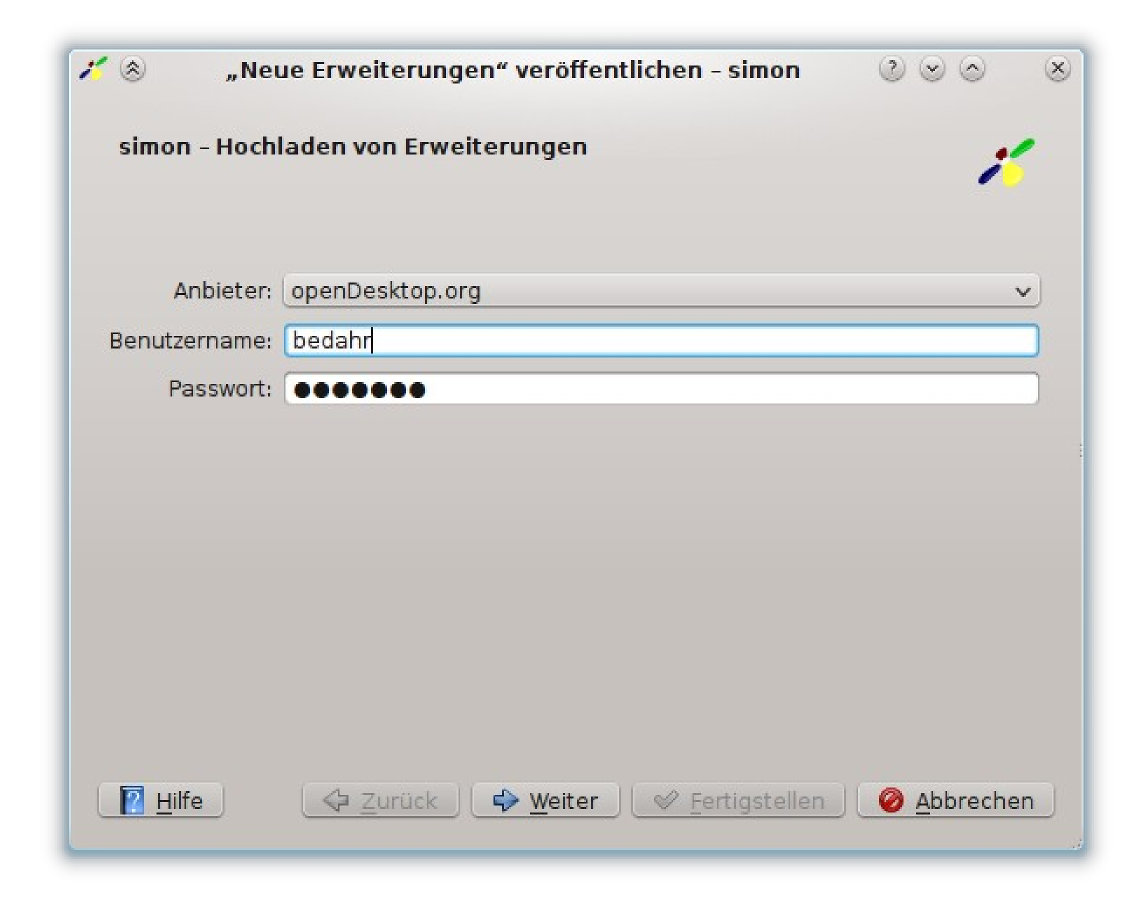
Szenario exportieren

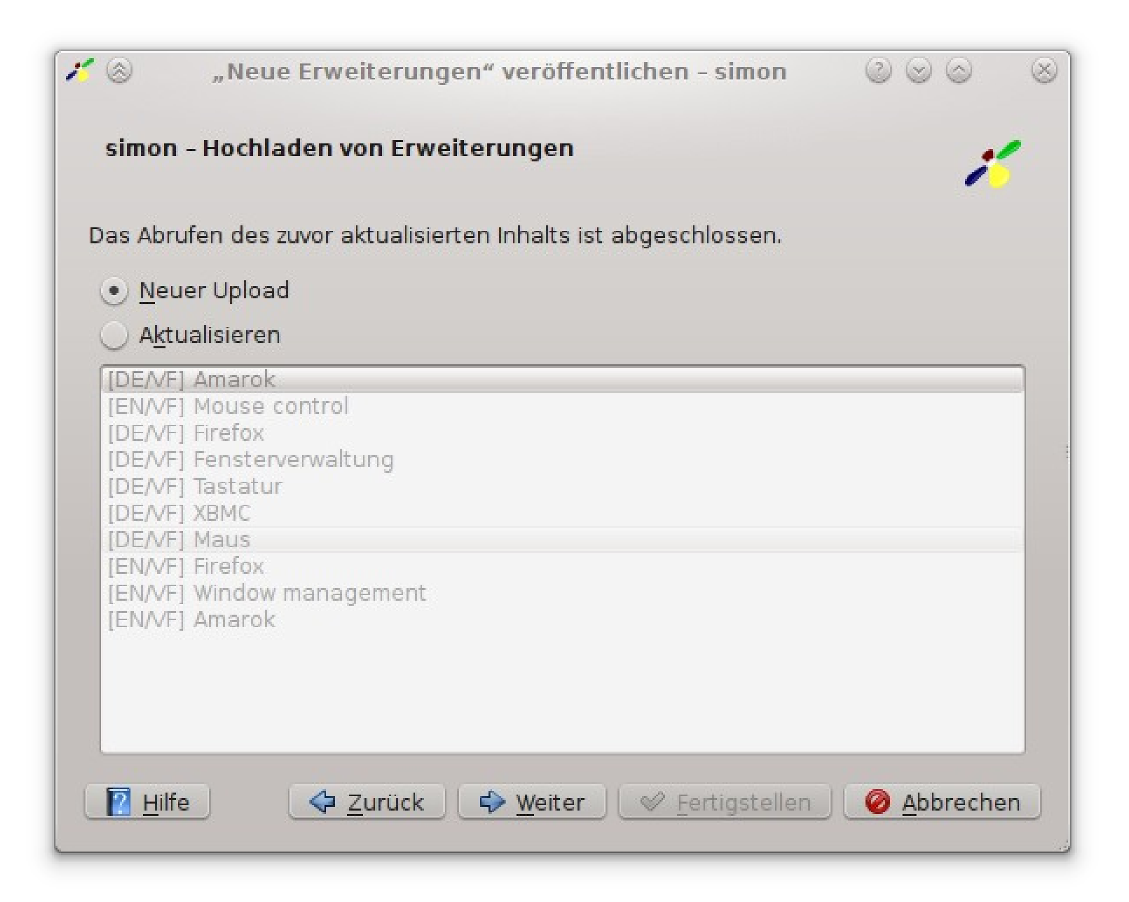
Szenarios können in eine lokale Datei (simons XML Szenariodateiformat) exportiert werden und diese

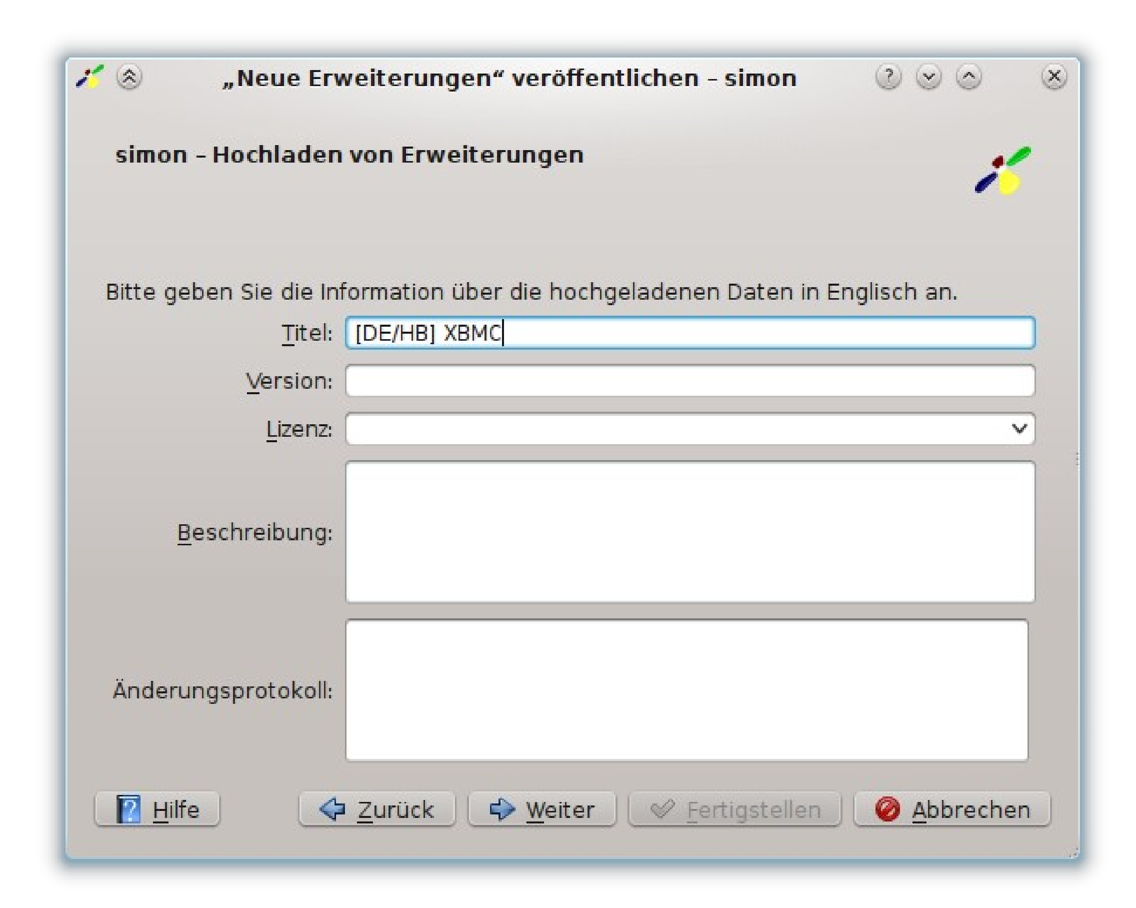
Datei dann in die simon Scenarios (http://kde-files.org/index.php?xcontentmode=692) Sektion der OpenDesktop Seite KDE-files.org (http://kde-files.org) hochgeladen werden.

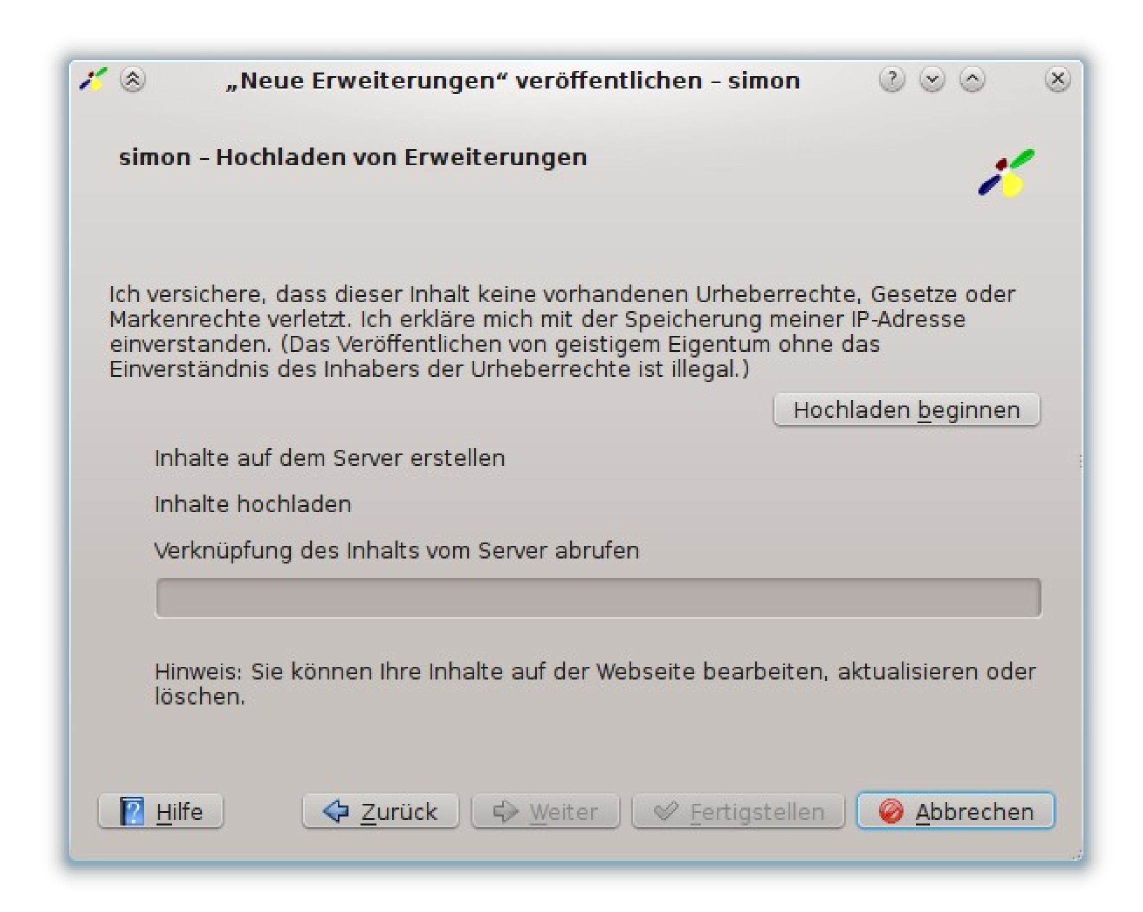
Um auf den OpenDesktop Seiten hochladen zu können, benötigen Sie einen Account auf der Seite. Die Registration (http://opendesktop.org/usermanager/new.php) ist allerdings sehr leicht und natürlich kostenlos.

simon erlaubt es Ihnen neue Inhalte direkt von der Benutzeroberfläche heraus hochzuladen (Export > Veröffentlichen).




„Neue Erweiterungen“ veröffentlichen – simon
simon – Hochladen von Erweiterungen
Das Abrufen des zuvor aktualisierten Inhalts ist abgeschlossen.
Neuer Upload
Aktualisieren
[DE/VF] Amarok
[EN/VF] Mouse control
[DE/VF] Firefox
[DE/VF] Fensterverwaltung
[DE/VF] Tastatur
[DE/VF] XBMC
[DE/VF] Maus
[EN/VF] Firefox
[EN/VF] Window management
[EN/VF] Amarok
Hilfe
Zurück
Weiter
Fertigstellen
Abbrechen


„Neue Erweiterungen“ veröffentlichen – simon
simon – Hochladen von Erweiterungen
Bitte geben Sie die Information über die hochgeladenen Daten in Englisch an.
Titel:
[DE/HB] XBMC
Version:
Lizenz:
Beschreibung:
Änderungsprotokoll:
Hilfe
Zurück
Weiter
Fertigstellen
Abbrechen

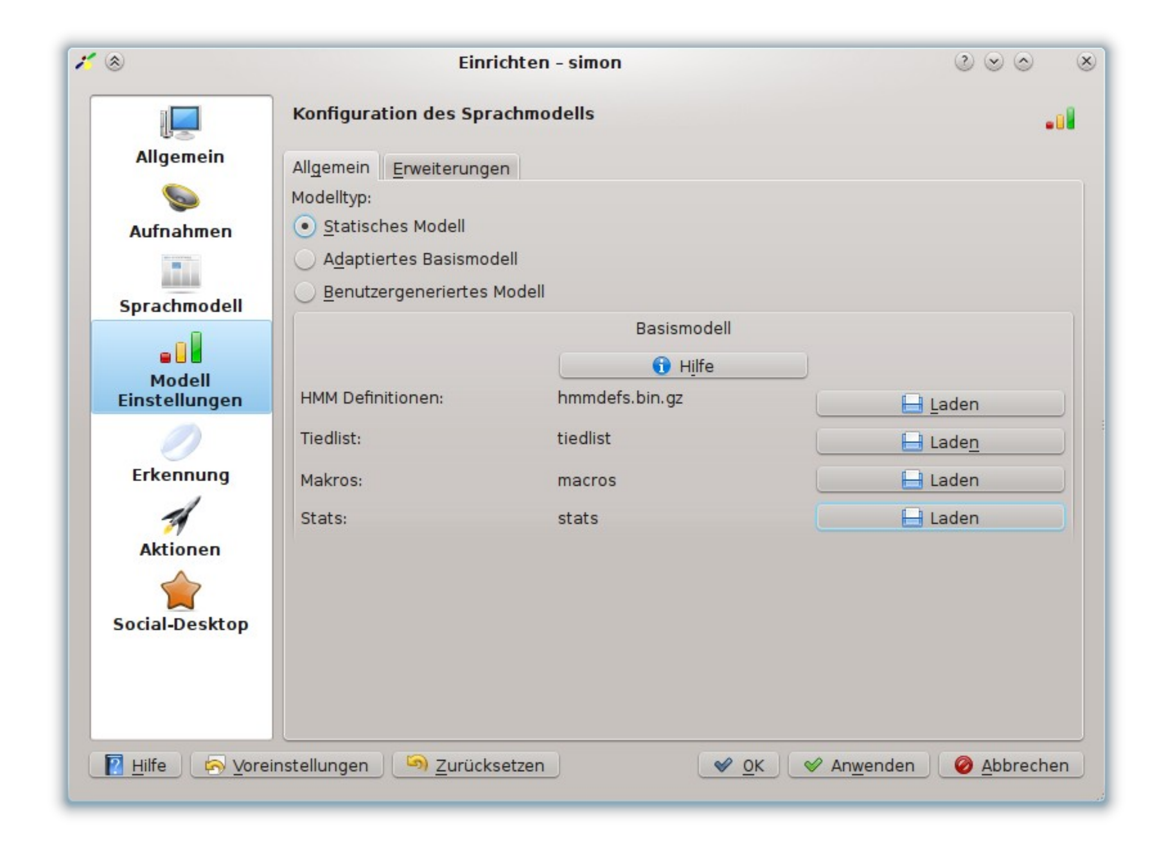


Um diese Funktionalität zu verwenden geben Sie einfach Ihre Accountdetails in die Soziale Desktop Konfiguration in der simon Konfiguration ein.

Basismodelle

Diese Sektion beschreibt wie Sie Basismodelle mit simon verwenden. Für generelle Informationen über Basismodelle, sehen Sie bitte das Hintergrund Kapitel.

Um simon zu konfigurieren Basismodelle zu verwenden, geben Sie einfach gültige Basismodelldateien in simons Konfiguration an: Einstellungen > simon einrichten... > Modelleinstellungen



Laden Sie die Dateien mit den entsprechenden "Laden" Knöpfen neben den Dateien die Sie laden wollen. Die Dateien werden zu einem internen Pfad kopiert sodass die Quelle entfernt werden kann nachdem das Basismodell übernommen wurde.

Für statische Modelle benötigen Sie weder "macros" noch "stats" Dateien aber simon wird die Kompilierung (die auch fuer statische Modelle notwendig ist um das Sprachmodell zu erzeugen) nicht ohne sie starten. Falls Ihr Basismodell diese Dateien nicht mitliefert können Sie einfach leeren Dateien angeben.

Vokabular

Das Vokabular ermöglicht es dem User, die verfügbaren Wörter zu verwalten.

Allgemein

Das Vokabular definiert, welche Wörter der Erkennungsprozess erkennen sollte. Jedes Wort, dass simon erkennen können soll muss in der Wortliste vorkommen.

Ein einzelner Eintrag in der Wortliste definiert exakt ein "Wort". Ein "Wort" bedeutet in simon eine einzigartige Kombination aus:

- Wordname

  (Das geschriebene Wort selbst)

- Terminal

(Grammatik-Kategorie; Z.B.: "Nomen", "Verb", etc.)

- Aussprache

  (Wie das Wort ausgesprochen wird; simon akzeptiert alle Formen von Lautschrift solange Sie weder Zahlen noch Sonderzeichen beinhaltet)

Das bedeutet, dass Mehrzahlformen oder Unterschiede in Groß- / Kleinschreibung als unterschiedliche Wörter behandelt werden. Dadurch erlaubt eine ausgefeilte Grammatik noch mehr Kontrolle.

Generell ist es empfohlen das Vokabular so klein wie möglich zu halten. Je mehr Wörter die Wortliste umfasst, desto wahrscheinlicher ist es, dass simon etwas falsches erkennt.

Beispiel Vokabular (bitte beachten Sie, dass die Terminale auf Nomen / Verb gesetzt sind um das Verstädnis zu erleichtern; Bitte sehen Sie die Grammatik Sektion warum dies nicht optimal ist):

Table 4-1. Beispiels-Vokabular

Wort	Terminal	Aussprache
Computer	Nomen	k O m p j u: t @ r
Internet	Nomen	I n t @ r n E t
Mail	Nomen	m E I l
schließen	Verb	S l i: s @ n

Aktives Vokabular

Das Vokabular, dass für die Erkennung verwendet wird, wird als aktives Vokabular bezeichnet.

Schattenlexikon

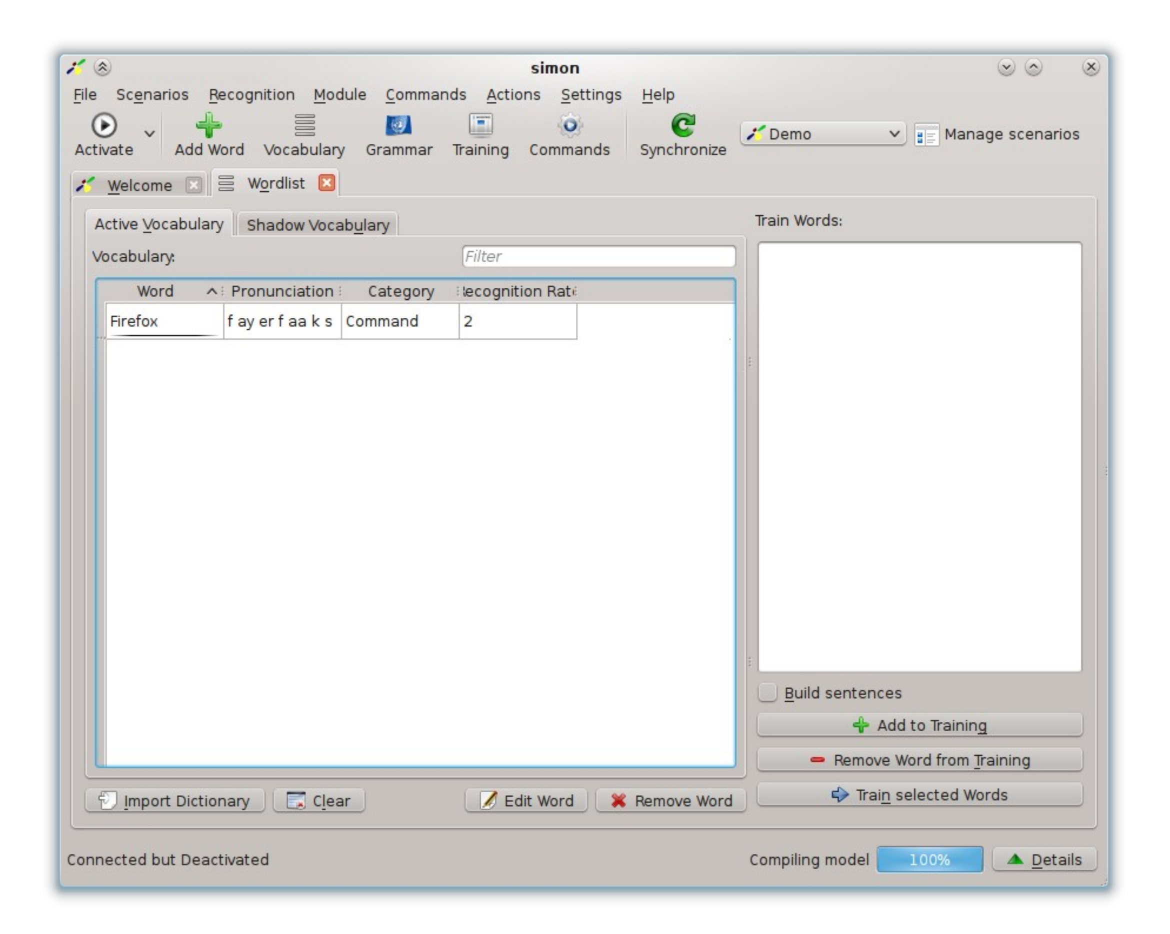
Wie bereits gesagt, sollte der Benutzer sein Vokabular so klein wie möglich zu halten. Anderseits sollte es Informationen über die Aussprache von möglichst vielen Wörter beinhalten um das hinzufügen neuer Wörter zu erleichtern.

simon bietet deshalb ein Referenzwörterbuch das "Schattenlexikon" genannt wird. Das Schattenlexikon ist ein zweites Vokabular, welches nicht für die Erkennung verwendet wird. Es kann mit fertigen Wörterbüchern gefüllt werden. Der Benutzer kann dann beim hinzufügen von neuen Wörtern von den Informationen aus dem Schattenlexikon profitieren.

Weil simon eine internationale Lösung ist, wird simon nicht mit einem Schattenlexikon ausgeliefert. Allerdings ist es sehr einfach Lexika über den Wörterbuch Importieren Assistenten zu importieren.

Verwalten des Vokabulars

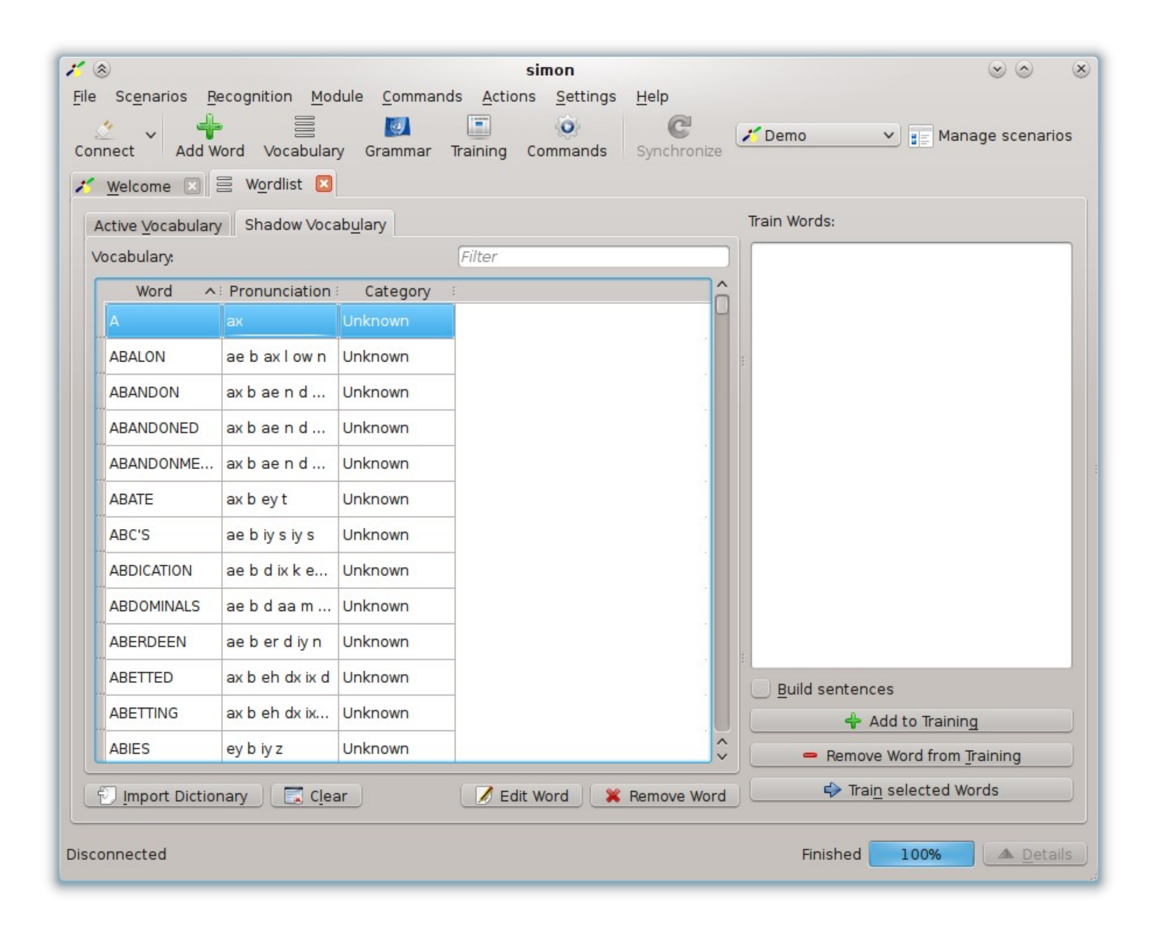
simon bietet ein "Vokabular" Menü welches das aktuelle Vokabular anzeigt.



Standardmäßig wird das aktive Vokabular angezeigt. Um das Schattenvokabular anzuzeigen wählen Sie die Registerkarte Schattenvokabular.

Für jedes Wort wird eine "Erkennungsrate" angezeigt. Die Erkennungsrate zeigt an wie oft ein Wort aufgenommen wurde (alleine oder zusammen mit anderen Wörtern).

Wenn diese Nummer nur Null oder Eins ist, wird der Eintrag in der Wortliste rot markiert (1: Hellrot; 0: Dunkelrot). Das ist eine Warnung. Wenn ein Wort eine Phonemkombination beinhaltet, welche nie trainiert wurde (Erkennungrate 0) kann das Modell *nicht* kompiliert werden. Allerdings wird simon einen entsprechende Fehlermeldung anzeigen sollte dieses Problem eintreten. Allgemein ist es eine gute Idee jedes Wort mindestens ein oder zweimal aufzunehmen (am besten währen das Wort hinzugefügt wird) um diesen Problem grundsätzlich aus dem Weg zu gehen.



Wörter hinzufügen

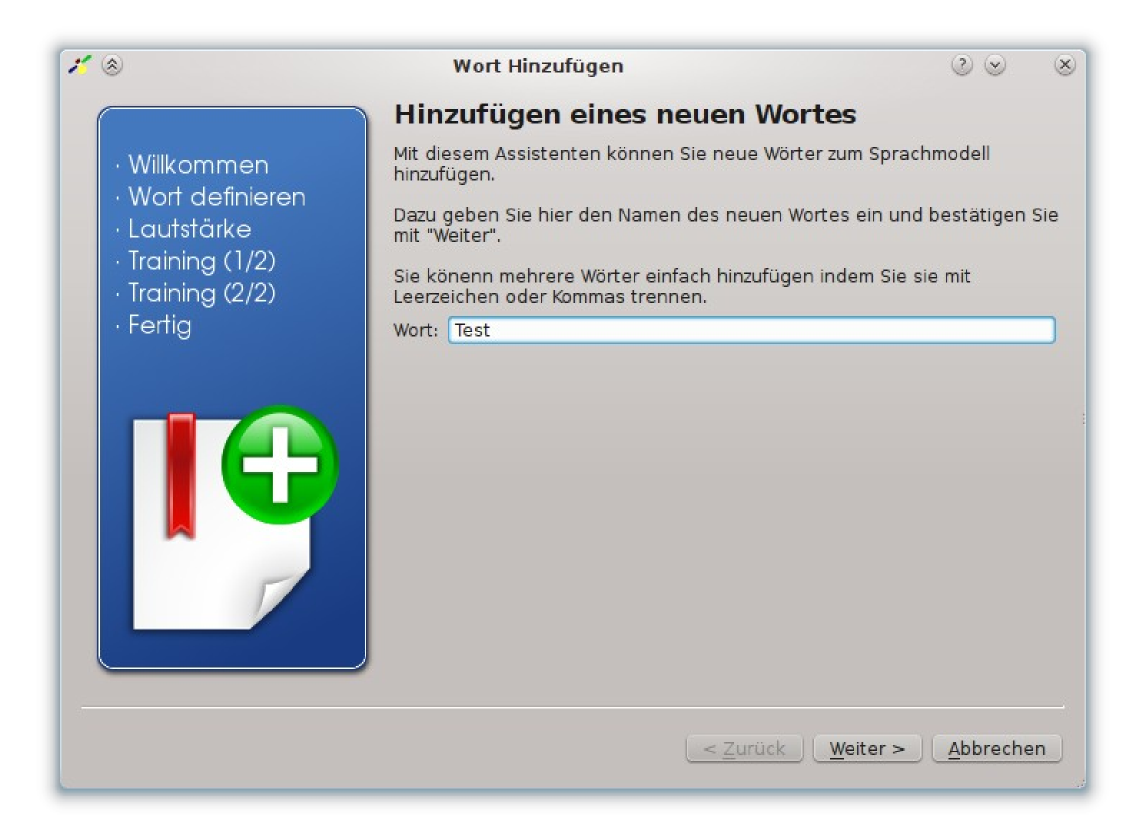
Um neue Wörter zum aktiven Vokabular hinzuzufügen, benutzen Sie bitte den Wort Hinzufügen Assistenten.

Wörter hinzuzufügen besteht im Prinzip aus zwei Schritten:

- Das Wort definieren
- Initiales Training

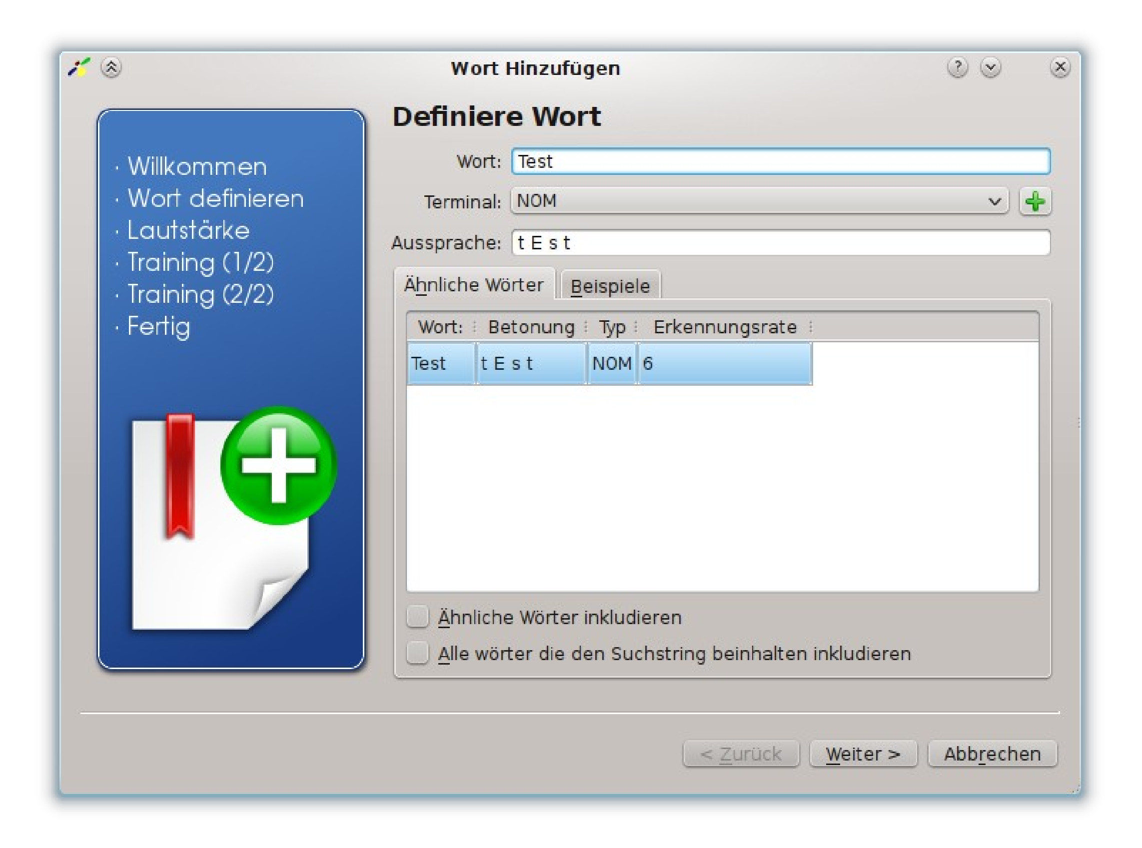
Definition des Wortes

Zuerst wird der Benutzer gefragt, welches Wort hinzugefügt werden soll.



Wenn der Benutzer fortfährt wird simon automatisch versuchen so viel Information wie möglich über das hinzuzufügende Wort im Schattenlexikon zu finden.

Wenn das Wort im Schattenlexikon vorkommt wird simon automatisch alle benötigten Felder ausfüllen (Terminal und Aussprache).



Alle Vorschläge vom Schattenlexikon sind in der Tabelle "Ähnliche Wörter" aufgelistet. Per Default werden nur exakte Treffer angezeigt. Allerdings kann dies geädert werden, wenn das Häkchen bei der Option "Zeige verwandte Wörter" gesetzt wird. Da die meisten Lautschriften relativ leicht zu lesen sind, kann dann leicht aus ähnlichen Wörtern die Korrekte Aussprache abgeleitet werden.

Das hägt natürlich von Ihrem Schattenlexikon ab. Wenn das Schattenlexikon die Wörter die Sie hinzufügen wollen nicht beinhaltet, müssen die notwendigen Felder manuell ausgefüllt werden.

Manche Wörterbücher die von simon importiert werden können (SPHINX, HTK) unterscheiden nicht zwischen Groß- und Kleinschreibung und beinhalten alle Wörter in Großbuchstaben. Natürlich kann die Schreibweise beim Hinzufügen frei geädert werden.

Einige Wörterbücher die von simon importiert werden können (SPHINX, PLS und HTK) beinhalten keine grammatikalische Information. Da alle Wörter in simon einen Terminal besitzen müssen, werden sie dem Terminal "Unbekannt" zugewiesen wenn Sie importiert werden. Sie sollten beim Hinzufügen der Wörter zum aktiven Lexikon dann den Terminal entsprechende ädern.

Manuell einen Terminal auswählen

Der Terminal des Wortes ist definiert als die grammatikalische Kategorie des Wortes. Das kann im Sinne der deutschen Grammatik "Nomen", "Verb" oder etwas komplett anderes wie erfundene Kategorien wie "Kommando" sein. Für mehr Information sehen Sie bitte die Grammatik Sektion.

Die Liste beinhaltet alle Terminale die bis jetzt in Ihrem aktiven und Schattenlexikon sowie in Ihrer Grammatik verwendet wurden.

Sie können neue Terminale zu der Liste hinzufügen indem Sie das grüne Plus verwenden.

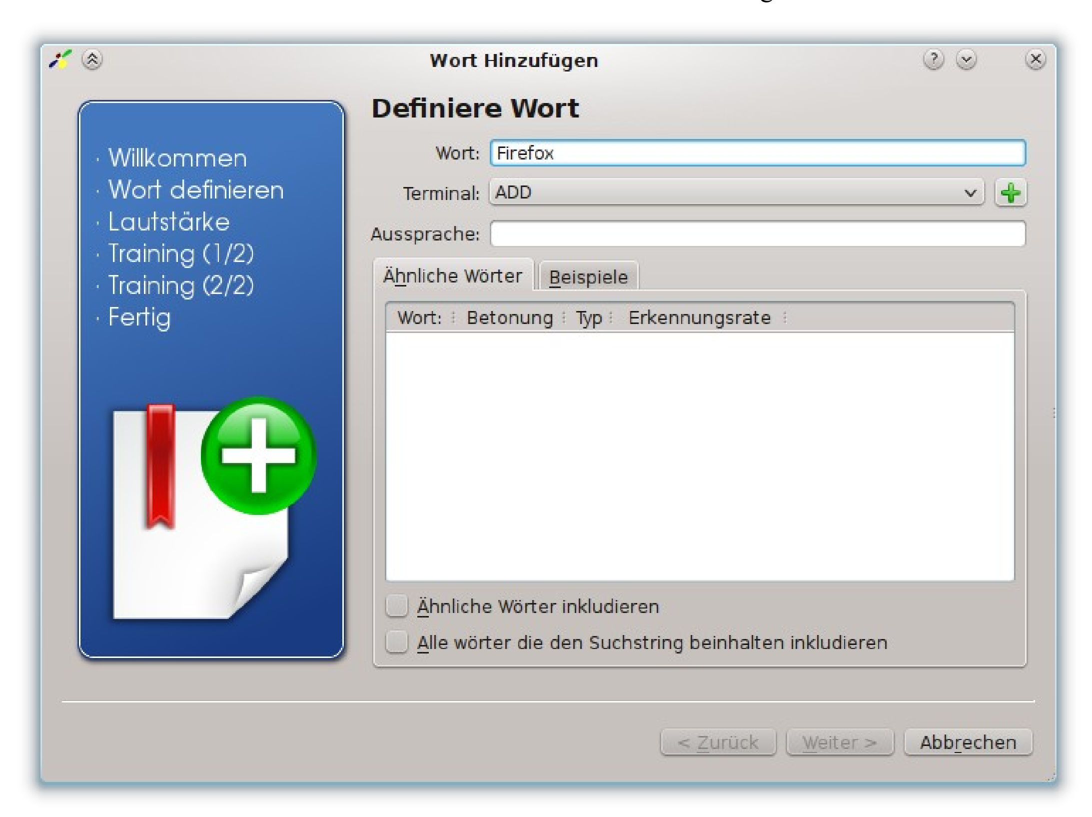
Phonetische Transkription manuell angeben

Die Aussprache ist etwas komplizierter. simon benötigt zwar keine bestimmten Lautschriftstandard solange er nur aus Buchstaben (keine Zahlen und Sonderzeichen) besteht. Allerdings ist es empfehlenswert die Lautschrift und das Phonemset des Schattenlexikons zu übernehmen.

Wenn Sie nicht wissen wie ein Wort transkripiert wird, können Sie das die Informationen aus dem Schattenlexikon verwenden - sogar wenn das Wort dort nicht vorkommt. Wenn wir zum Beispiel das Wort "Firefox" hinzufügen wollen um den Webbrowser Firefox zu starten ist dieses Wort höchstwahrscheinlich nicht in unserem Schattenlexikon vorhanden.

(Ich verwende das HADIFIX-BOMP als mein Schattenlexikon)

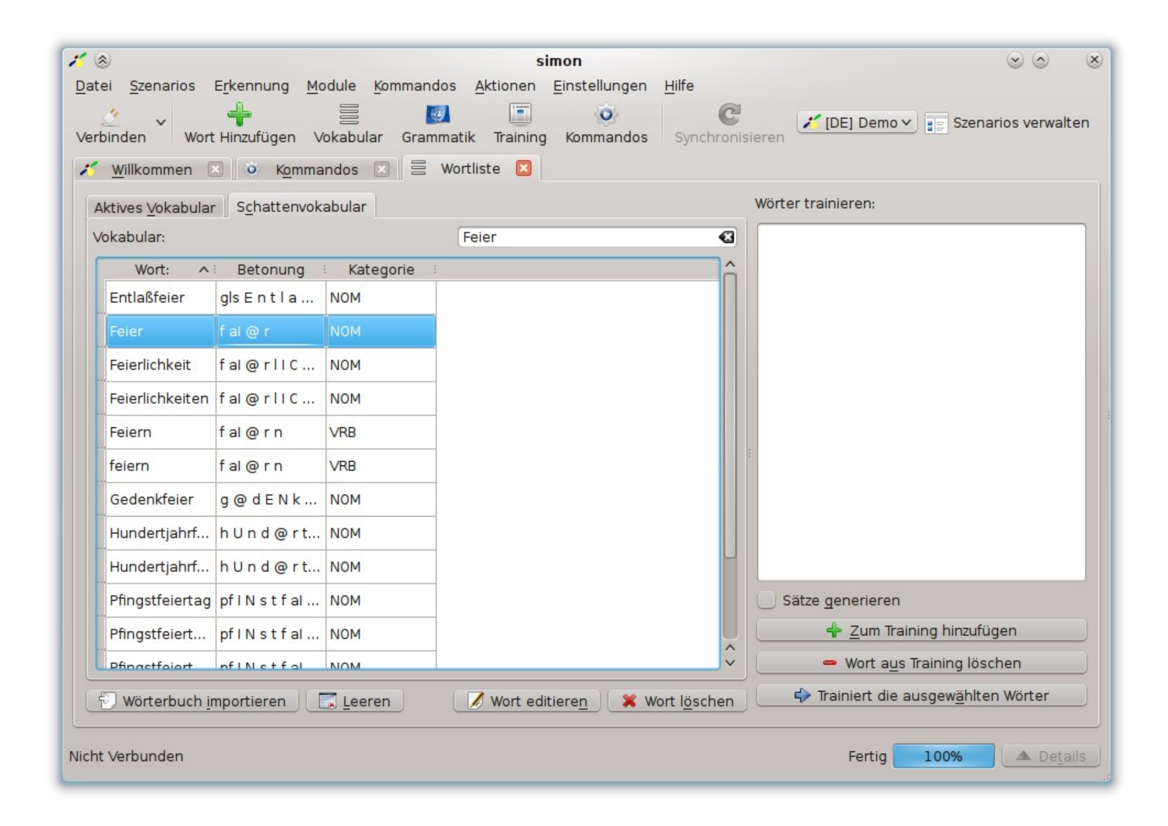
"Firefox" ist nicht in unserem Schattenlexikon enthalten. Deshalb schlägt simon keine Werte vor.



Wir wissen aber, dass das Wort Firefox gleich ausgesprochen wird wie die Wörter "Feier" (Fest) und "Fox" (Tanz) welche beide im BOMP Lexikon vorkommen. Öffnen wir also das Vokabular (Sie können den Assistenten offen lassen) indem Sie "Vokabular" über die simon Toolbar auswählen.

Wechseln Sie zum Schattenvokabulr indem Sie auf die Registerkarte Schattenvokabular klicken.

Suchen wir mit der "Filter"-Box über der Liste nach "Feier":



Wir können sehen, dass das Wort "Feier" mit "f aI @ r" transkripiert wurde. Filtern wir jetzt nach "Fox" statt "Feier" und wir sehen, dass "Fox" "f O k s" ausgesprochen wird. Daraus können wir ohne Probleme ableiten, dass das Englische "Firefox" wie ein Deutsches "Feierfox" also "f aI @ r f O k s" ausgesprochen und transkripiert wird.

So können Sie die Betonung von ähnlich ausgesprochenen Wörtern ableiten. Diese Strategie hat auch den Vorteil, dass selbst Laien mit einem guten Schattenlexikon sehr gute Transkriptionen finden können und dabei auch die gleichen Phoneme verwenden die im Schattenlexikon vorkommen und somit automatisch für Wörter verwendet werden, die im Schattenlexikon gefunden wurden und nur ins aktive Lexikon übernommen worden sind.

Wir können jetzt die korrekte Aussprache angeben und den Terminal ädern.



Trainieren des Wortes

Um den Assistenten abzuschließen können Sie jetzt das Wort zwei mal aufnehmen. Falls Sie das nicht tun wollen oder z.B. ein statisches Modell verwenden können Sie die folgenden zwei Seiten überspringen.

Weil Sie jetzt gleich zwei Trainingsaufnahmen aufnehmen werden wird simon die Lautstärkeeinpegelung anzeigen um sicherzustellen, dass die Lautstärke optimal eingestellt ist. Für mehr Informationen sehen Sie bitte die Lautstärkeeinpegelung Sektion.

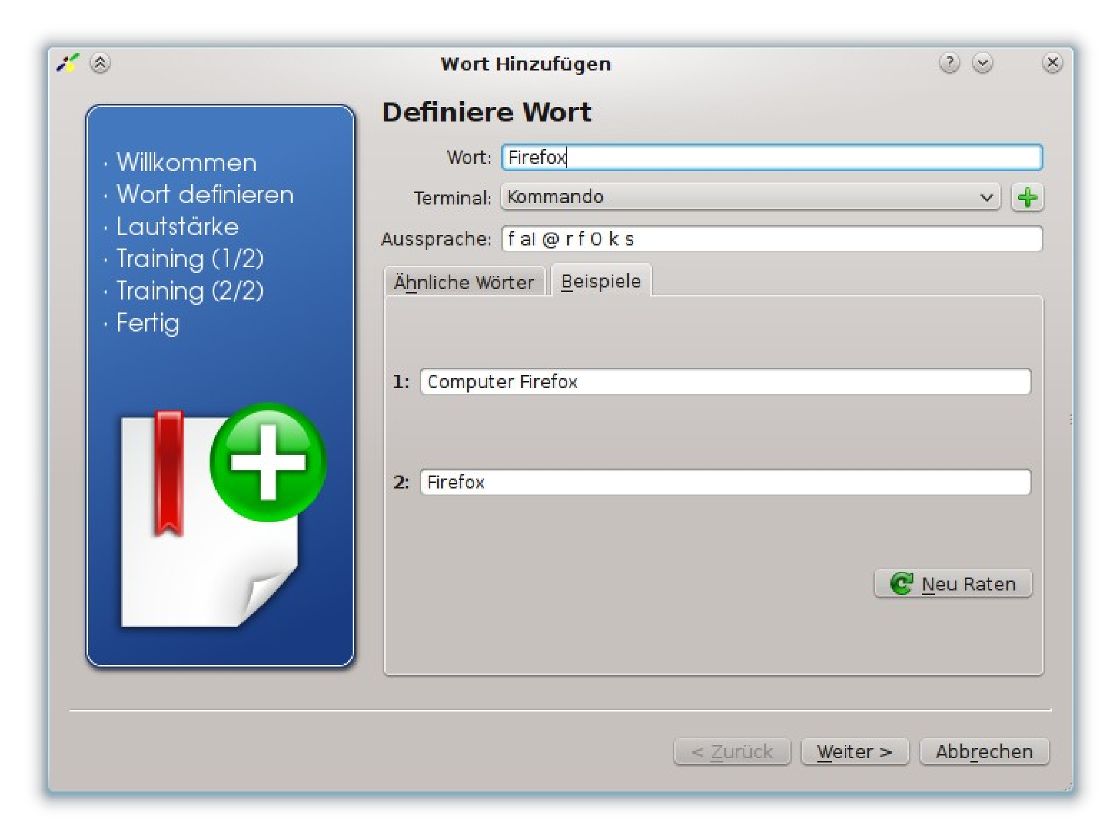
simon wird automatisch versuchen Sie beim Training nach Sätzen statt einzelnen Wörtern zu fragen. Dazu wird simon Ihre Grammatik durchforsten nach Satzstrukturen die den Terminal des neuen Wortes beinhalten und dann anhand der Terminale der anderen Wörter in ihrem aktiven Lexikon einen gültigen Satz bilden.

Zum Beispiel: Sie haben eine Grammatikstruktur "Trigger Kommando" und haben das Wort "Computer" vom Terminal "Trigger" in Ihrem Vokabular. Sie fügen dann ein Wort "Firefox" vom Terminal "Kommando" hinzu. simon wird dann automatisch den Satz "Computer Firefox" bilden und diesen für das erste Training des Wortes verwenden weil es laut der Grammatik ein gültiger Satz ist.

Wenn simon keinen gültigen Satz bilden kann (z.B. wenn der Terminal (noch) nicht in ihrer Grammatik vorkommt oder Sie keine Grammatik definiert haben), wird simon Sie einfach nur nach dem einzelnen Wort fragen.

simon stellt sicher, dass die Beispiele für das erste Training korrekt sind, aber Sie können die Vorschläge jederzeit ädern. Wechseln Sie dazu zum "Beispiel" Reiter in der "Definiere Wort" Seite des

Assistenten.



Sie können die Beispiele frei editieren. Sie können sogar Wörter verwenden die noch nicht in Ihrem aktivem Lexikon vorkommen solange Sie sie hinzufügen bevor Sie das nächste mal das Modell synchronisieren obwohl das nicht empfohlen ist.

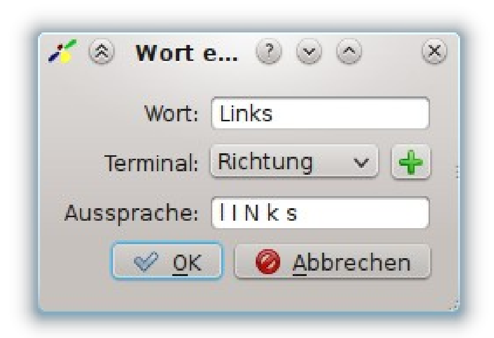
Nun bleibt nur noch das Wort aufzunehmen.



Stellen Sie sicher den Richtlinien der Aufnahme Sektion zu folgen.

Ein Wort ädern

Um ein Wort zu editieren, wählen Sie es im angezeigten Vokabular aus und klicken Sie auf Wort editieren.



Hier können Sie Namen, Terminal und Aussprache des ausgewählten Wortes ädern.

Ein Wort entfernen

Um ein Wort aus dem Sprachmodell zu entfernen, wählen Sie es im Vokabular aus und wählen Sie "Ausgewähltes Wort löschen".



Wie Sie sehen, bietet der Dialog vier Auswahlmöglichkeiten:

- Das Wort zum Terminal "Unbenutzt" verschieben.

  Weil Sie (hoffentlich) den Terminal "Unbenutzt" in Ihrer Grammatik nicht verwenden, werden Wörter die so markiert sind nicht läger für die Erkennung verwenden. Sie werden vom aktiven Vokabular entfernt bevor das Modell erstellt wird weil keine Grammatikstruktur sie mehr referenziert.

  Wenn Sie den Terminal "Unbenutzt" in Ihrer Grammatik verwenden wollen können Sie natürlich einen anderen Terminal für unbenutzte Wörter verwenden. Sie können den Terminal über den Wort editieren Dialog setzen.

  Um das Wort erneut zu verwenden, setzen Sie es einfach weder auf den richtigen Terminal. Assoziierte Trainingsdaten bleiben erhalten.

- Entferne das Wort vom aktiven Vokabular

  Diese Option wird das ausgewählte Wort vom aktiven Lexikon (und damit von der Erkennung)

entfernen aber wird das Wort im Schattenvokabular weiterführen(sollte das Schattenlexikon das selbe Wort bereits beinhalten, wird es nicht noch einmal eingefügt). Alle Aufnahmen (Trainingsdaten) die das Wort beinhalten bleiben erhalten.

Um das Wort wieder zu verwenden, fügen Sie es wieder zum aktiven Vokabular dazu. Wenn Sie ein neues Wort mit dem selben Namen hinzufügen werden die Werte dieses Wortes vorgeschlagen werden. Es gehen also keine Daten unwiederbringlich verloren.

- Lösche das Wort aber behalte assoziierte Aufnahmen

  Entfernt das Wort komplett aber behält Aufnahmen in denen es vorkommt. Wenn Sie ein weiteres Wort mit dem selben Namen hinzufügen werden die Aufnahmen wieder damit assoziiert.

  Seien Sie vorsichtig mit dieser Option da das neu hinzugefügte Wort möglicherweise anders transkripiert wird und dieser Unterschied nicht automatisch erkannt und beachtet ewrden kann (simon wird dann bei der Modellgenerierung die neue Aussprache auch bei den alten Aufnahmen erwarten).

  Wählen Sie diese Option nicht wenn die Aufnahmen für das Wort falsch (betont) waren.

- Das Wort komplett löschen

  Entfernt das Wort. Alle Aufnahmen die das Wort beinhalten werden ebenfalls entfernt.

  Diese Option lässt keine Spuren von weder dem Wort selbst noch assoziierten Sprachdaten auf dem System zurück.

  Da Aufnahmen global sind (nicht zu einem spezifischen Szenario zugewiesen) werden auch Aufnahmen die in Trainingssitzungen von anderen Szenarios aufgenommen wurden entfernt wenn Sie dieses Wort beinhalten. Verwenden Sie diese Funktion vorsichtig.

Spezialtraining

Bitte sehen Sie die Sektion Spezialtraining im Trainingskapitel.

Wörterbuch Importieren

simon bietet die Möglichkeit große Wörterbücher als Nachschlagewerk zu importieren. Dieses Nachschlagewerk wird Schattenlexikon genannt.

Wenn der Benutzer ein neues Wort hinzufügt muss er die folgenden Charakteristika des Wortes angeben:

- Wortname
- Terminal
- Phonetische Transkription

Diese Charakteristika werden automatisch aus dem Schattenlexikon genommen wenn das Schattenlexikon ein Wort mit dem selben Wortnamen führt. Ein großes, hochqualitatives Schattenlexikon kann deshalb sehr hilfreich sein beim Hinzufügen neuer Wörter weil der Benutzer sich in den meisten Fällen kaum mit den Details wie dem Phonemset auseinander setzen muss.



Seit Version 0.3 können Sie Wörterbücher auch direkt in das aktive Vokabular importieren. Diese Option ist hauptsächlich nützlich um es leichter zu machen von selbst entwickelten manuellen Lösungen auf simon umzusteigern und um Modelle aus älteren simon Versionen zu importieren (z.B. aus Version 0.2). Sie werden *fast nie* ein sehr großes Wörterbuch in das aktive Vokabular importieren wollen.

Sie finden eine Liste verfügbarer Wörterbücher die mit simon getestet wurden auf dem simon Wiki (http://www.cyber-byte.at/wiki/index.php/Deutsch:_Schattenw%C3%B6rterbuch).

simon kann derzeit mit fünf verschiedenen Wörterbuchformaten umgehen:

- HADIFIX
- HTK
- PLS
- SPHINX
- Julius

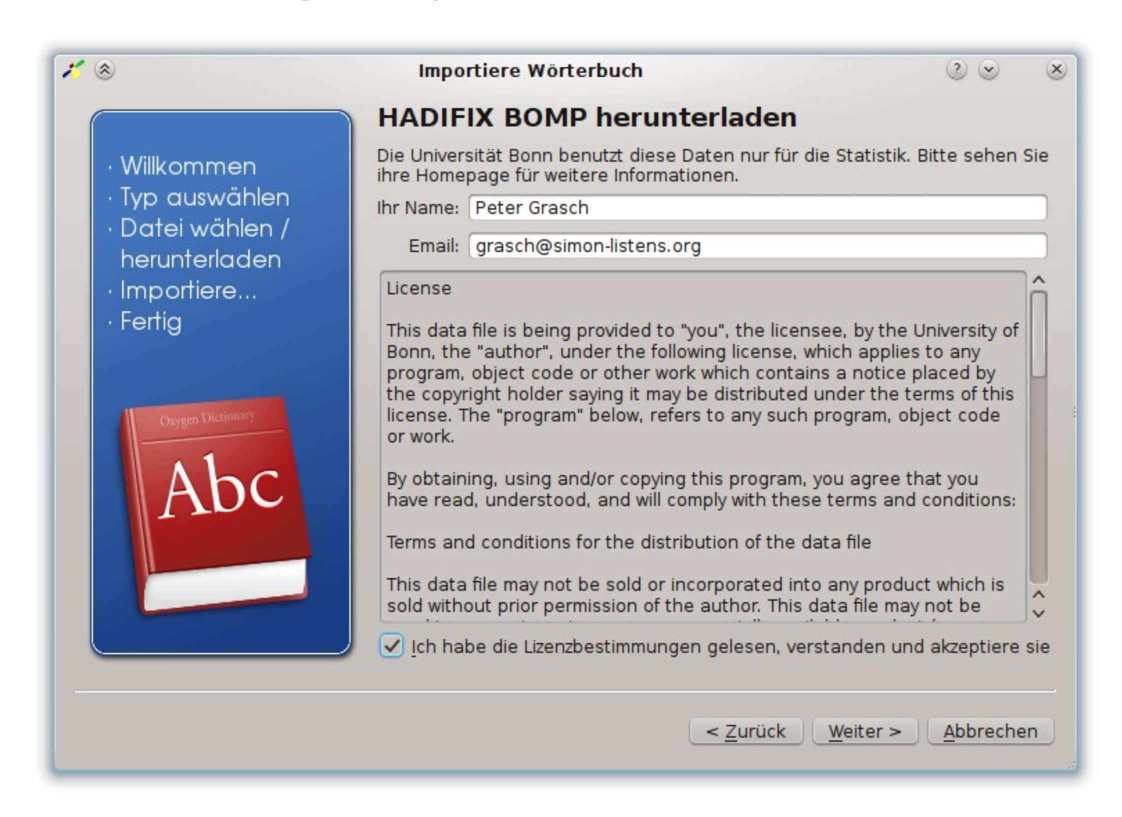
HADIFIX Wörterbuch

simon kann HADIFIX Wörterbücher importieren

Ein Beispiel eines HADIFIX Wörterbuches ist das Deutsche HADIFIX BOMP (http://www.sk.uni-bonn.de/forschung/phonetik/sprachsynthese/bomp).

Hadifix Wörterbücher bieten sowohl Terminal- als auch Ausspracheinformationen.

Durch eine spezielle Ausnahme in Ihrer LIzenz ist das Simon Listens Team stolz darauf Ihnen einen direkten Download und Import des ausgezeichneten HADIFIX BOMPs in simon anbieten zu können.



Mit dem automatischen BOMP Import können Sie, nachdem Sie Namen und e-Mail Adresse für die Universität Bonn angegeben haben, das Wörterbuch direkt von einem simon listens Server importieren.

HTK Wörterbuch

simon kann HTK Wörterbücher importieren

Ein Beispiel eines HTK Wörterbuchs ist das Englische Voxforge Wörterbuch (http://www.repository.voxforge1.org/downloads/SpeechCorpus/Trunk/Lexicon/).

HTK Wörterbücher bieten Ausspracheinformationen aber keine Terminale. Alle Wörter werden dem Terminal "Unbekannt" zugewiesen.

PLS Wörterbuch

simon kann PLS Wörterbücher importieren.

Ein Beispiel eines PLS Wörterbuches ist das Deutsche GPL Wörterbuch von Voxforge (http://www.repository.voxforge1.org/downloads/de/Trunk/Lexicon/).

PLS Wörterbücher bieten Ausspracheinformationen aber keine Terminale. Alle Wörter werden dem Terminal "Unbekannt" zugewiesen.

SPHINX Wörterbuch

simon kann SPHINX Wörterbücher importieren.

Ein Beispiel eines SPHINX Wörterbuches ist das Wörterbuch für Mexikanisches Spanisch (http://speech.mty.itesm.mx/~jnolazco/proyectos.htm).

SPHINX Wörterbücher bieten Ausspracheinformationen aber keine Terminale. Alle Wörter werden dem Terminal "Unbekannt" zugewiesen.

Julius Wörterbuch

simon kann Julius Wörterbücher importieren.

Ein Beispiel eines Julius Vokabulars sind Wortlisten von simon 0.2.

Julius Wörterbücher bieten Ausspracheinformationen und Terminalinformationen.

Grammatik

Die Grammatik definiert, welche Wortkombinationen korrekt sind.

Allgemein

Beispiel: Sie wollen simon benutzen um Programme zu starten und die Programmfenster zu schließen wenn Sie fertig sind. Sie würden gerne die folgenden Kmmandos verwenden:

- "Computer, Internet" um einen Browser zu öffnen
- "Computer, Mail"

  Um einen Mail Client zu öffnen
- "Computer, schließen"

  Um das aktuelle Fenster zu schließen.

Wenn wir der deutschen Grammatik folgen, würde die Grammatik die folgenden Einträge beinhalten:

Table 4-2. Beispiels-Vokabular

Wort	Terminal
Computer	Nomen
Internet	Nomen
Mail	Nomen
schließen	Verb

Um die oben definierten Sätze zu erlauben würde wir folgende Grammatik verwenden:

- "Nomen Nomen" für Sätze wie "Computer Internet"
- "Nomen Verb" für Sätze wie "Computer schließen"

Das würde zwar funktionieren, aber würde auch Kombinationen wie "Computer Computer", "Internet Computer", "Internet Internet", etc. erlauben die alle nicht wirklich sinnvoll sind. Um die Erkennungsgenauigkeit zu erhöhen, können wir unsere Grammatik optimieren.

Es ist wichtig im Hinterkopf zu halten, dass jeder Benutzer mit simon seine eigene "Sprache" definiert. Das bedeutet, dass Sie in keiner Weise an die Grammatikregeln der zugrunde liegenden "echten" Sprache (Deutsch, Englisch, etc.) gebunden sind. Für simple Steuerung des Computers ist es sogar empfehlenswert eigene Grammatikregeln zu erfinden um grammatikalische Unterschiede die für den Anwendungsfall nicht wichtig sind zu eliminieren und unsinnige Sätze zu verbieten.

In dem Beispiel oben ist es beispielsweise nicht relevant, dass "schließen" ein Verb ist oder das "Computer" und "Internet" Nomen sind. Stattdessen können wir sie als etwas deklarieren das besser reflektiert was sie in unserem Szenario sein sollen:

Table 4-3. Verbessertes Beispielvokabular

Wort	Terminal
Computer	Trigger
Internet	Kommando
Mail	Kommando
schließen	Kommando

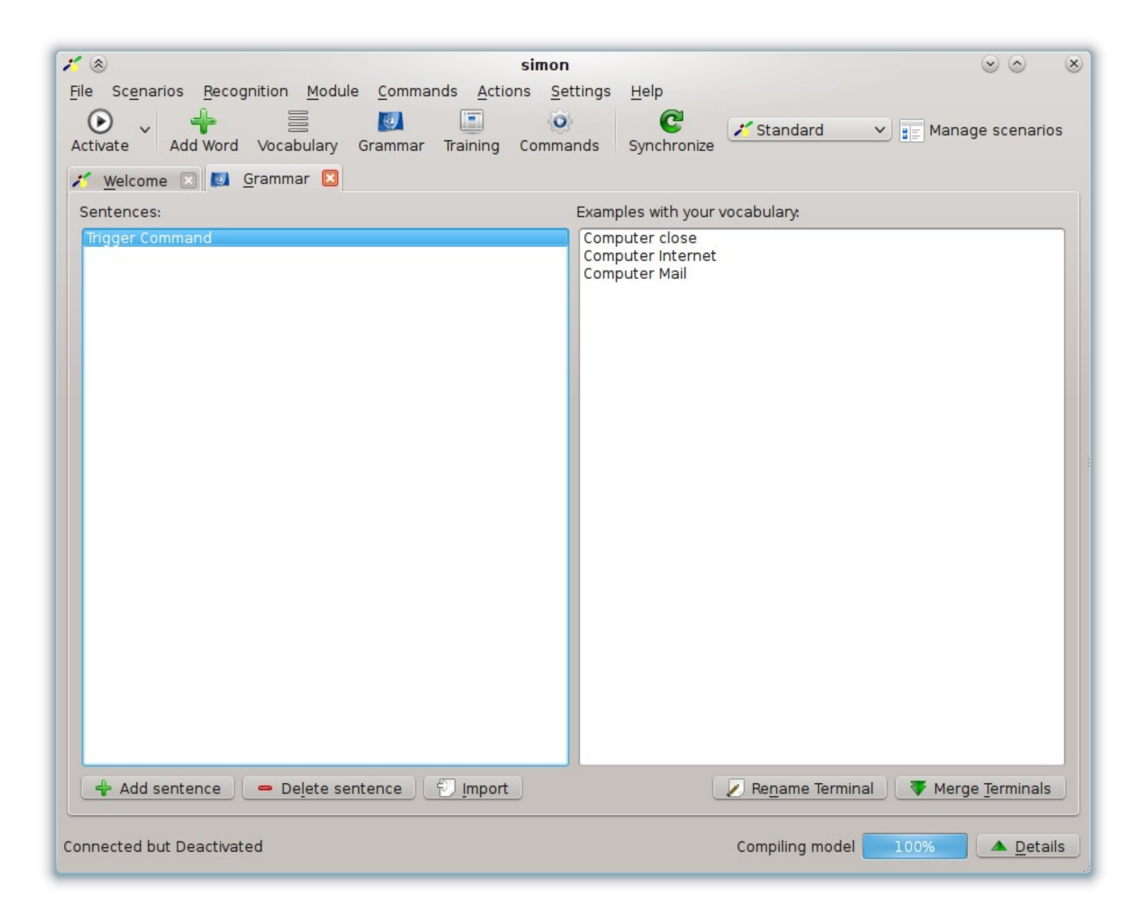
Jetzt ädern wir die Grammatik zu folgendem:

- "Trigger Kommando"

Dies erlaubt simon alle gewünschten Kombinationen zu erkennen. Aber es limitiert die Erkennugsergebnisse auch auf exakt diese drei möglichen Sätze. Vor allem in größeren Modellen kann eine clevere Grammatik und ein dazupassendes Vokabular große Unterschiede in der Erkennungsrate bewirken.

Die Grammatik definieren

simon bietet eine einfache, Text basierte Oberfläche um die Grammatik zu definieren. Sie können einfach alle erlaubten Sätze (ohne Satzzeichen) wie oben beschrieben eingeben.



Wenn Sie einen Satz links auswählen wir das Feld rechts automatisch echte Sätze die mit Ihrem aktuellen Vokabular dadurch ermöglicht werden anzeigen.

Die Beispielsektion erlaubt maximal 35 Beispiele sodass die Liste nicht vollstädig sein muss.

Eine Grammatik importieren

Zusätzlich zum Eingeben der Grammatik Satz bei Satz kann simon auch Satzstrukturen von Textdateien "erlernen".



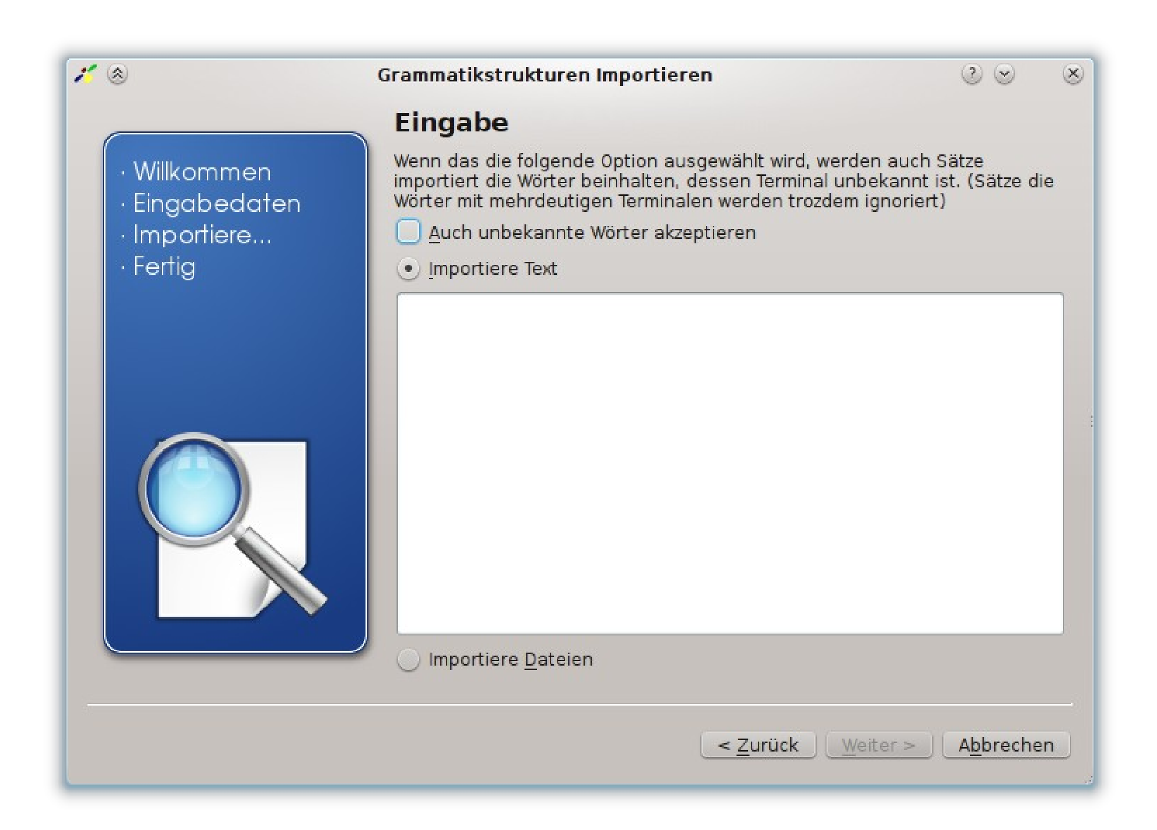
simon kann Textdateien einlesen aber erlaubt auch direkte Eingabe von Text in ein Eingabefeld.

Wenn wir bspw. ein Vokabular wie vorhin in der allgemeinen Sektion haben:

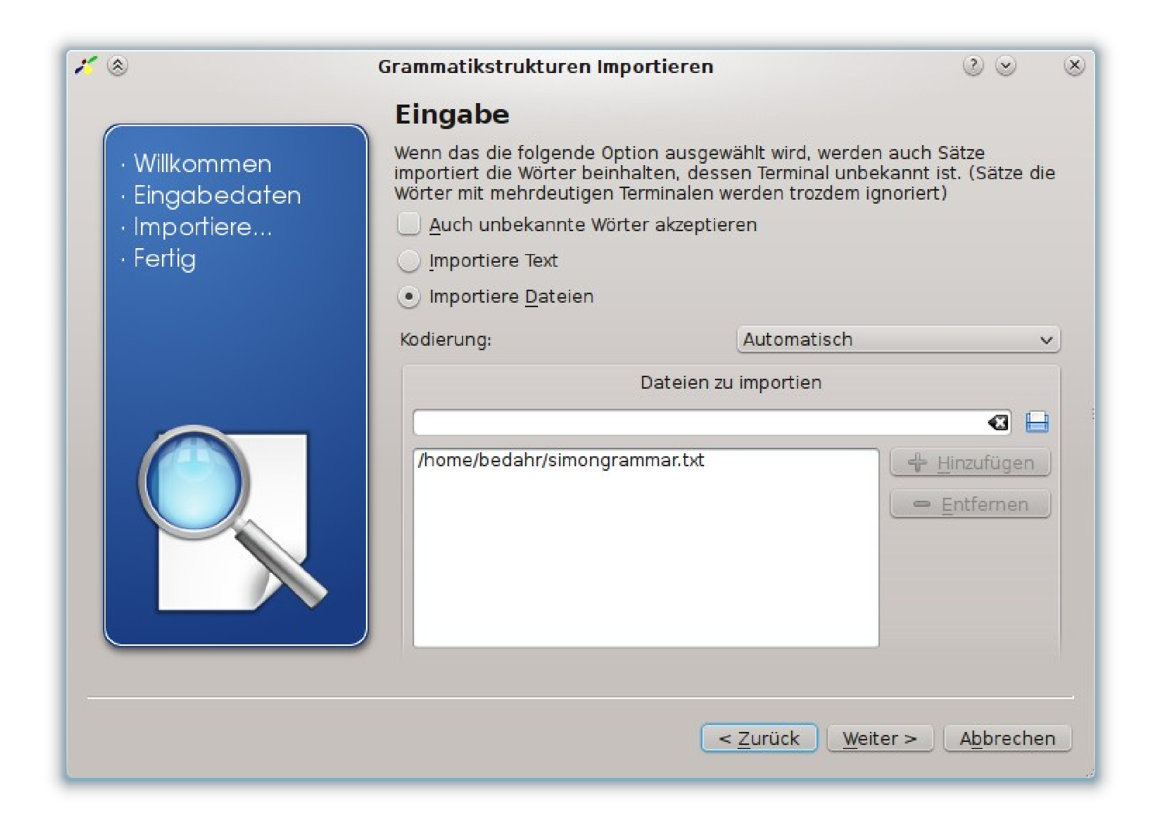
Table 4-4. Verbessertes Beispielvokabular

Wort	Terminal
Computer	Trigger
Internet	Kommando
Mail	Kommando
schließen	Kommando

Wir wollen also ein System haben, dass den Satz "Computer Internet!" erkennt. Also geben wir entweder den Text über Importiere Text ein oder erstellen eine Textdatei mit dem Inhalt "Computer Internet!" (jedes Satzzeichen würde funktionieren) und speichern es als "simongrammar.txt" um die Option Importiere Dateien zu verwenden.


Grammatikstrukturen Importieren
Willkommen
Eingabedaten
Importiere...
Fertig
Eingabe
Wenn das die folgende Option ausgewählt wird, werden auch Sätze importiert die Wörter beinhalten, dessen Terminal unbekannt ist. (Sätze die Wörter mit mehrdeutigen Terminalen werden trozdem ignoriert)
Auch unbekannte Wörter akzeptieren
Importiere Text
Importiere Dateien
< Zurück
Weiter >
Abbrechen


simongrammar.txt - KWrite
Datei
Bearbeiten
Ansicht
Extras
Einstellungen
Hilfe
Neu
Öffnen
Speichern
Speichern unter
Computer Internet!
Zeile: 1 Spalte: 19
EINF
ZEILE
simongrammar.txt



simon wird dann den eingegebenen Text bzw. all die angegebenen Textdateien einlesen (in diesem Fall nur die Datei "simongrammar.txt") und jedes einzelne Wort in sowohl dem aktiven als auch dem Schattenlexikon nachschlagen (die Definition im aktiven Lexikon hat Vorrang über eventuelle Definitionen im Schattenlexikon). Es wird dann das Wort mit dem entsprechenden Terminal ersetzten.

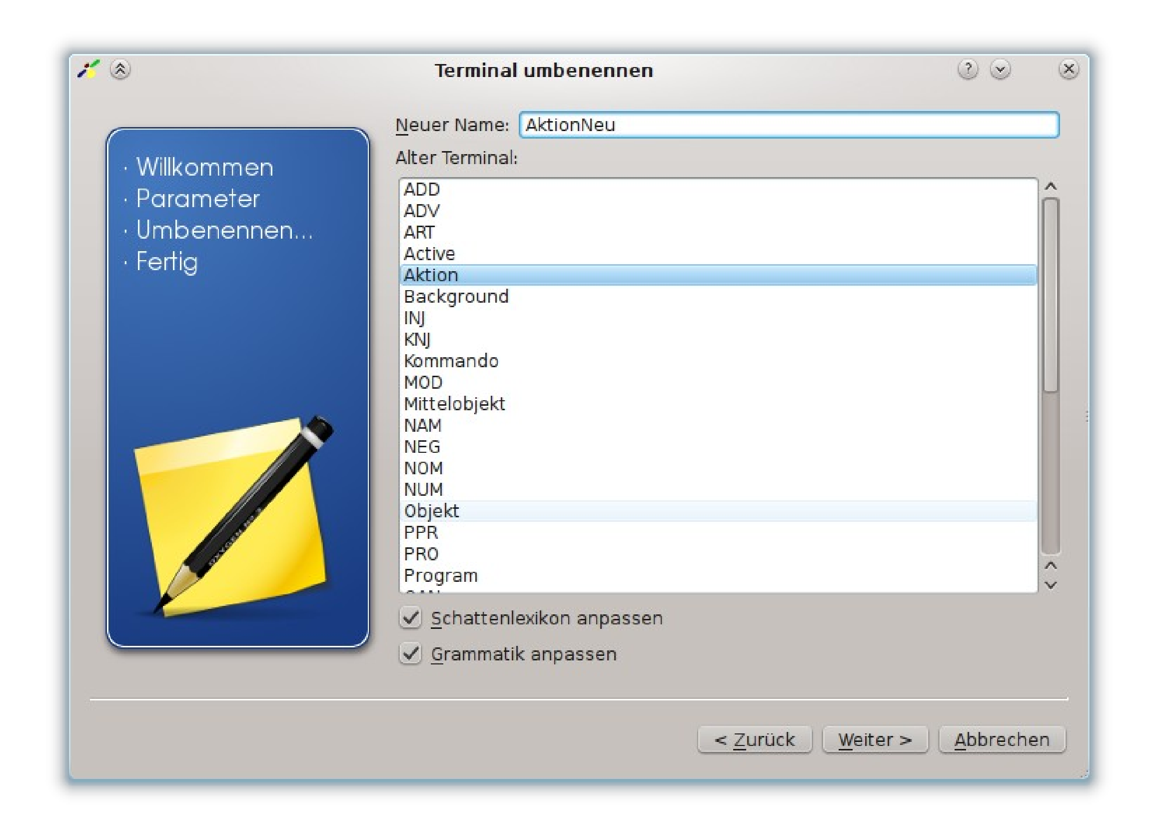
In unserem Beispiel würde das bedeuten, dass werden Satz "Computer Internet" finden würde. simon würde dann erkennen, dass "Computer" vom Terminal "Trigger" und "Internet" vom Terminal "Kommando" ist. Deswegen würde simon "lernen", dass "Trigger Kommando" offensichtlich ein gültiger Satz ist und ihn automatisch zur Grammatik hinzufügen.

Der Import segmentiert den Eingabetext automatisch nach Satzzeichen (".", ," - ", "!", etc.) also jeder natürliche Text sollte funktionieren. Der Import wird automatisch alle doppelt vorkommenden Satzstrukturen zusammenfügen (sogar über verschiedene Dateien verstreut) und wird alle möglichen grammatikalischen Satzstrukturen hinzufügen wenn Sätze Wörter verwenden die mit mehr als einem Terminal assoziiert sind.

Der Import wird alle Sätze ignorieren die ein oder mehrere Wörter beinhalten die im Sprachmodell nicht gefunden werden konnten außer die Option "Auch unbekannte Wörter akzeptieren" wurde aktiviert. In diesem Fall werden unbekannte Wörter mit dem generischen Terminal "Unbekannt" ersetzt.

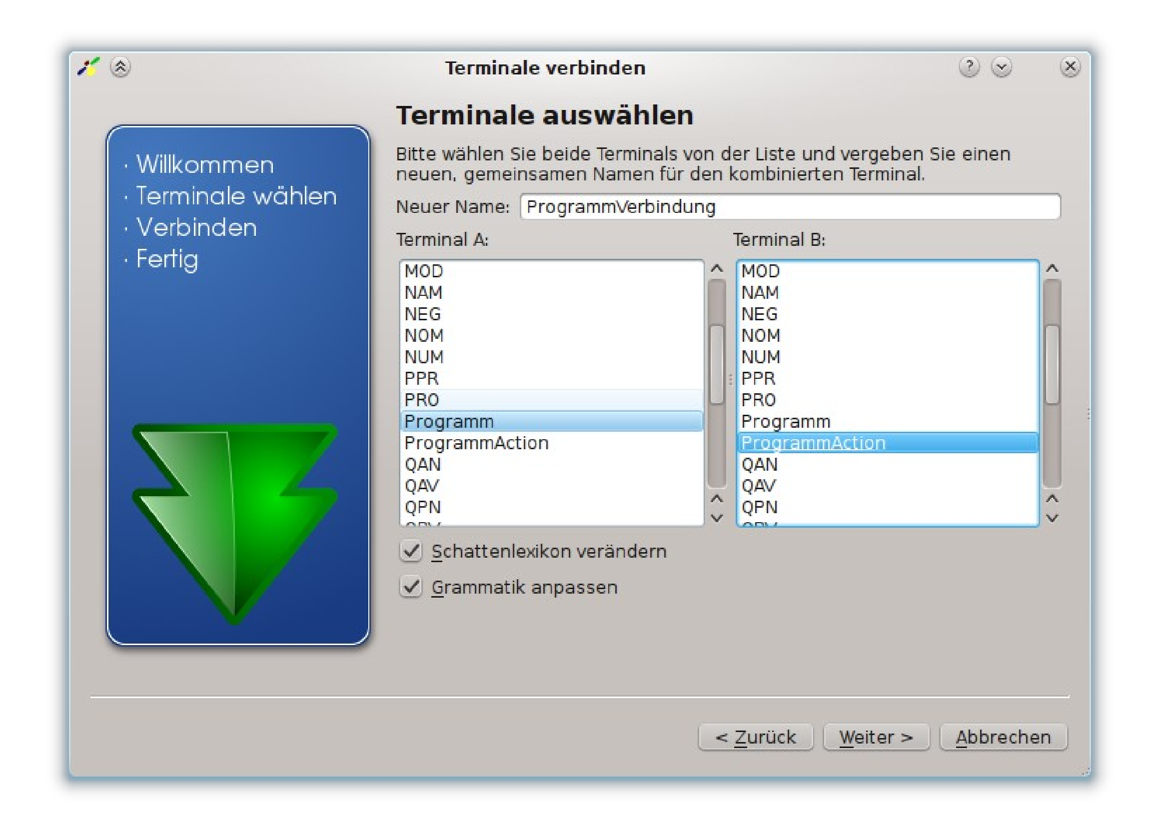
Terminale umbenennen

Der Terminal umbenennen Assistent erlaubt es dem Benutzer einzelne Terminale im aktiven Vokabular, dem Schattenlexikon und in der Grammatik umzubenennen.



Terminale verbinden

Der Terminal verbinden Assistent erlaubt es zwei Terminale in einen neuen zusammenzuführen. Er ädert dabei das aktive Lexikon und auf Wunsch auch das Schattenlexikon und die Grammatik.



Diese Funktion ist besonders nutzvoll wenn Sie Ihre Grammatikstrukturen vereinfachen wollen.

Training

Mit dem Trainingsmodul können Sie Ihr Akustikmodell und damit die Erkennungsrate verbessern.

Die Oberfläche listet alle installierten Trainingstexte in einer Tabelle mit drei Spalten:

- Name

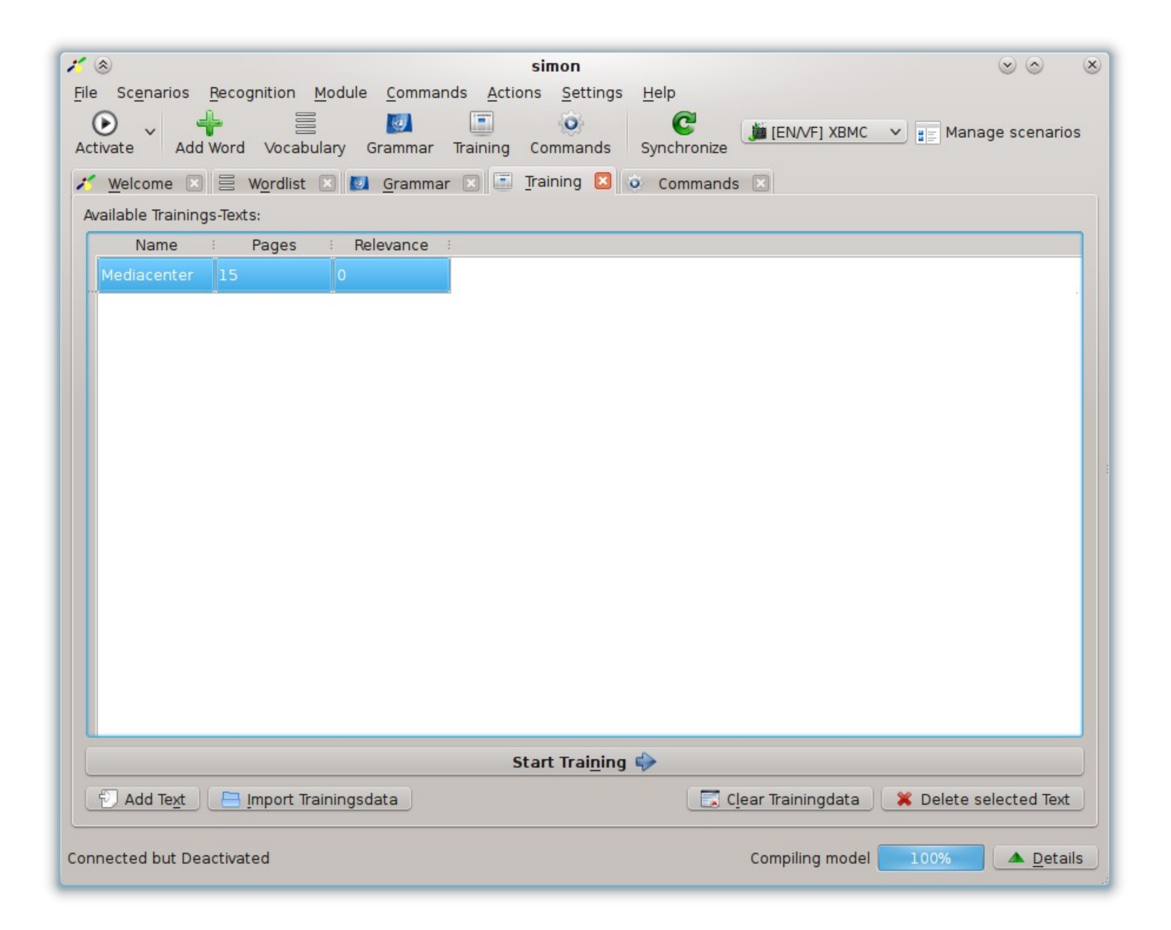
  Ein beschreibender Name für den Text.

- Seiten

  Die Anzahl von "Seiten" aus der der Text besteht. Jede Seite repräsentiert eine Aufnahme.

- Erkennungsrate

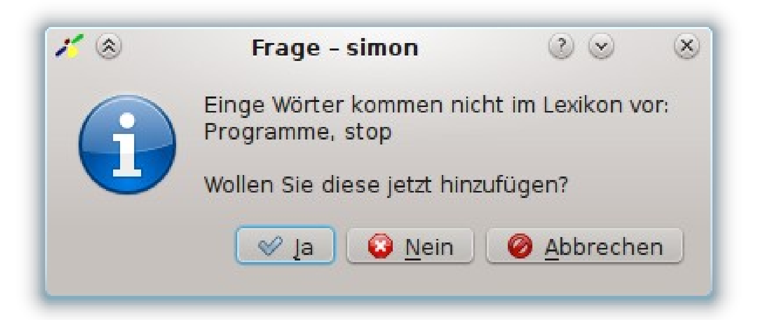
  Analog zum Vokabular; Zeigt an wie wahrscheinlich es ist, dass simon die Wörter die der Text beinhaltet korrekt erkennt (höher ist besser). Der Wert ist dabei einfach ein Mittelwert der Erkennungsraten aller Wörter im Text.



Um das Akustikmodell zu verbessern - und damit die Erkennungsrate müssen Sie Trainingstexte aufnehmen. Trainingsdaten bestehen im Allgemeinen aus zwei Teilen:

- Sprachaufnahmen Ihrer Stimme
- Transkriptionen dieser Aufnahmen

Das aktive Wörterbuch wird verwendet um die Wörter zu transkripieren (die Worte in ihre Phoneme aufschlüsseln) aus denen der Text besteht den Sie vorlesen (trainieren). Deshalb muss jedes Wort, dass in dem Text verwendet wird auch in Ihrem aktiven Vokabular vorkommen. Ist das nicht der Fall, wird simon Sie warnen und die Möglichkeit anbieten die fehlenden Wörter hinzuzufügen.

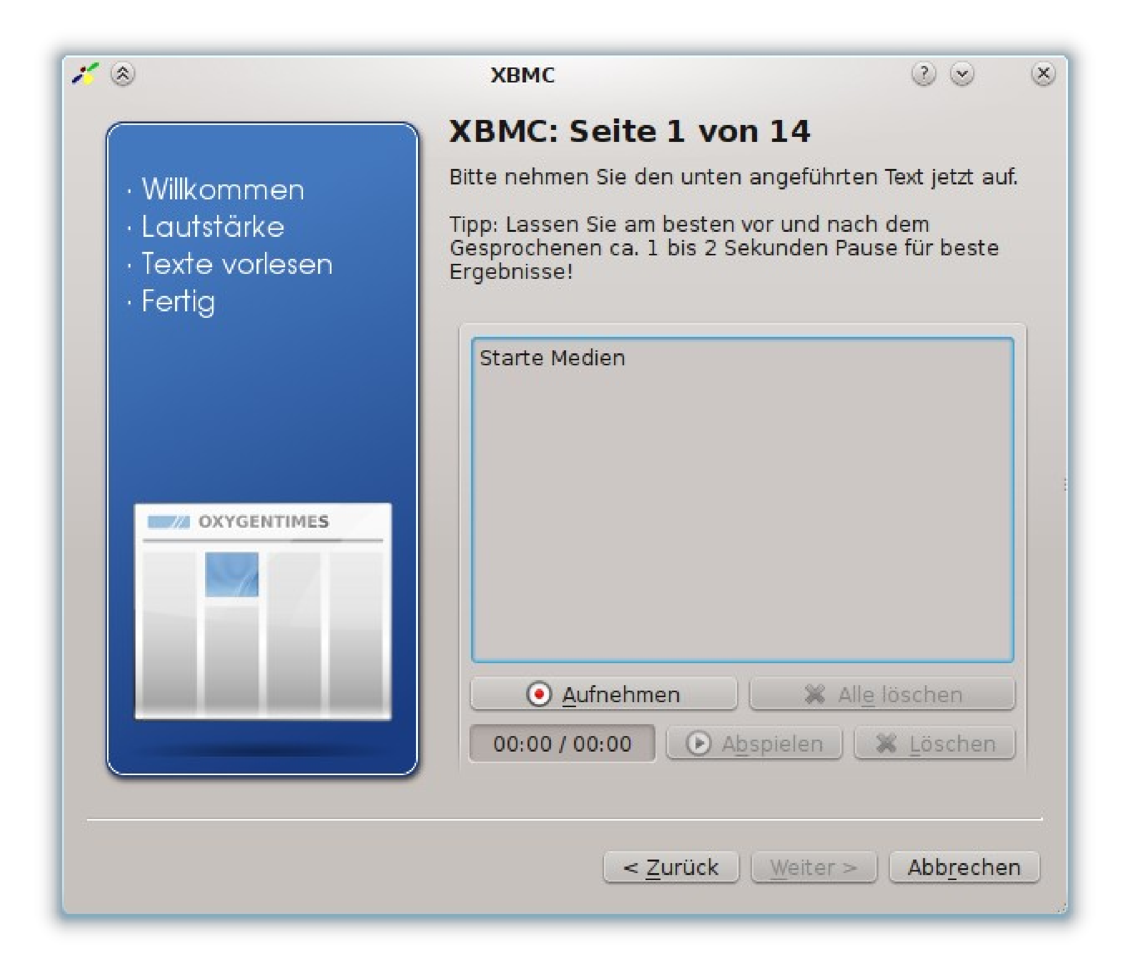


Das Hinzufügen ist ganz gleich als würden Sie ein neues Wort hinzufügen aber der Assistent wird

automatisch mehrmals durchlaufen um die Information über alle fehlende Wörter zu sammeln.Dieser Vorgang kann natürlich jederzeit abgebrochen werden. Sollten Sie das machen, wird simon Sie fragen ob er die bisher fertig definierten Wörter hinzufügen soll oder ob alle Änderungen rückgägig gemacht werden sollen. Wenn Sie alle Wörter hinzugefügt haben werden die Änderungen gespeichert und das Training des zuvor ausgewählten Textes startet automatisch.

Das Training (vorlesen) eines Trainingstexts funktioniert exakt gleich wie das erste Training beim Hinzufügen eines neuen Wortes.

Stellen Sie sicher den Richtlinien der Aufnahme Sektion zu folgen.



Speicherorte

Trainingstexte sind an zwei Stellen gespeichert:

- Linux: `~/.kde/share/apps/simon/texts`

  Windows: `%appdata%\.kde\share\apps\simon\texts`

  Die Texte des aktuellen Benutzers. Können mit simon gelöscht und hinzugefügt werden (siehe unten).

- Linux: ``` `kde4-config --prefix`/share/apps/simon/texts ```

  Windows: `(Installationsordner)\share\apps\simon\texts`

Systemweite Texte. Sie werden bei jedem (Betriebssytem-) Benutzer auf dem System angezeigt und können innerhalb von simon je nach Betriebssystemberechtigung höchstwahrscheinlich nicht gelöscht werden.

Dieser Ordner eignet sich besonders gut für Systemadministratoren die allen Benutzern ein gemeinsames Set an Trainingsdaten geben wollen.

Die XML Dateien (eine für jeden Text) können von einem Speicherort zum anderen verschoben werden. Abhägig von den zugrunde liegenden Betriebssystemberechtigungen wird dies aber wahrscheinlich Administratorenrechte benötigen.

Texte hinzufügen

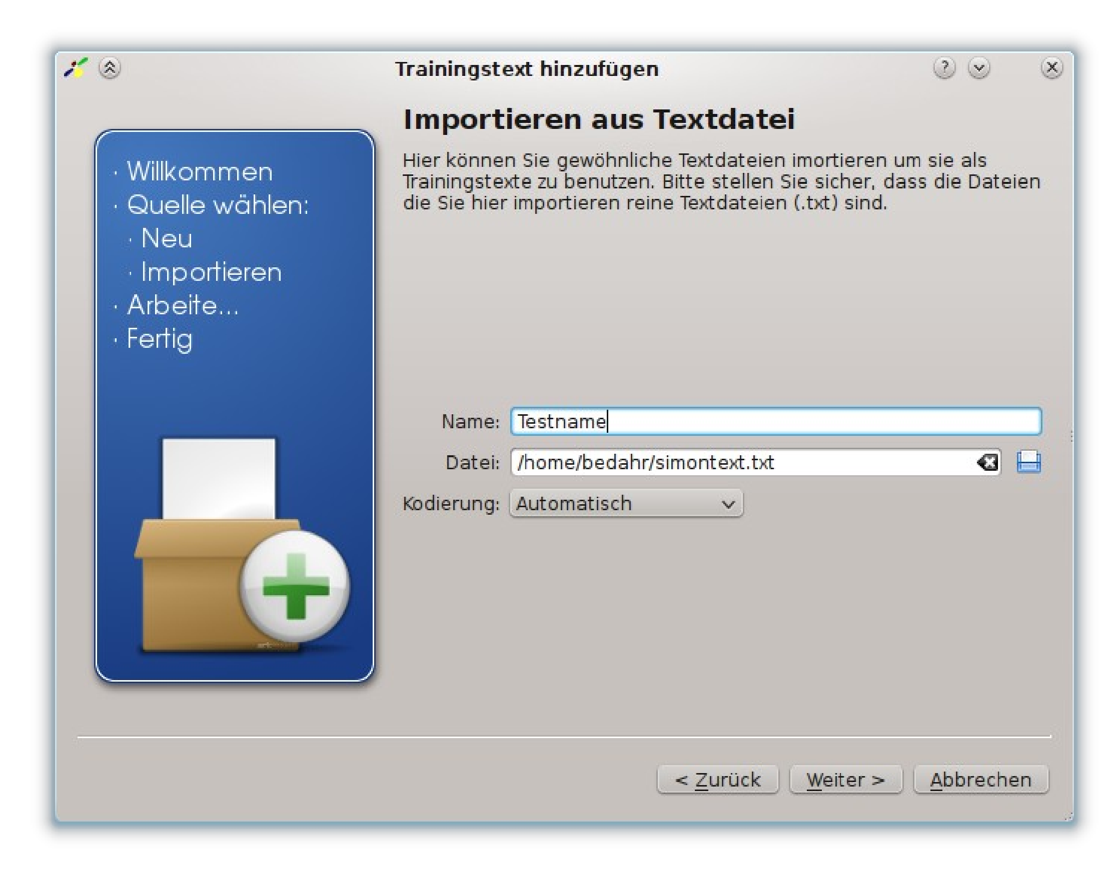


Der Trainingstexte hinzufügen Assistent bietet einen einfachen Weg neue Texte zu simon hinzuzufügen.

Wenn lokale Textdateien importiert werden, wird simon automatisch versuchen einzelne Sätze aufzuspalten um Sie beim Training eigenen "Seiten" (Aufnahmen) zuzuordnen. Der Algorithmus behandelt zwischen normalen Satzzeichen (".", "!", "?", "...", """,...) und Absätzen als "Sätze". Jeder "Satz" wird im Trainingstext auf einer Seite und somit auf einer eigenen Aufnahme sein.

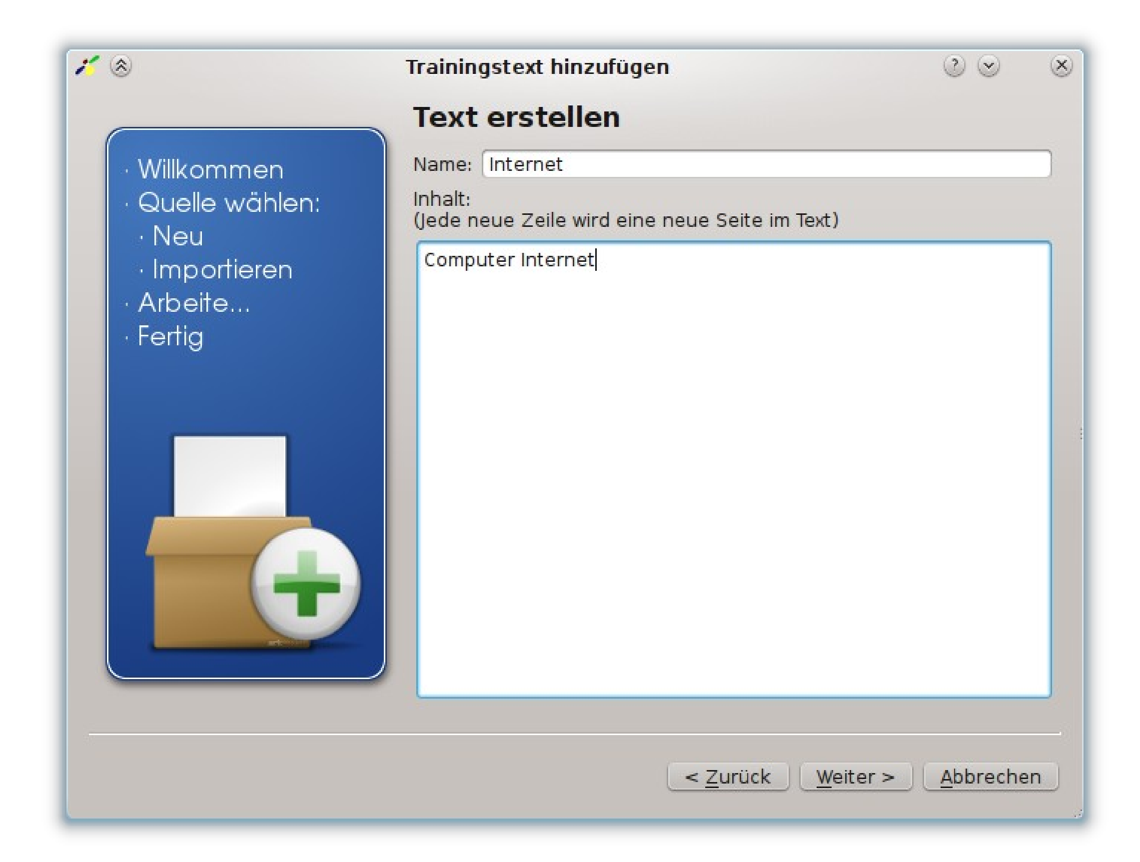
simon unterstützt zwei Quellen für neue Trainingstexte:

Trainingstexte hinzufügen Assistent



Tragen Sie den Trainingstext einfach in das entsprechende Eingabefeld ein.

Lokale Textdateien

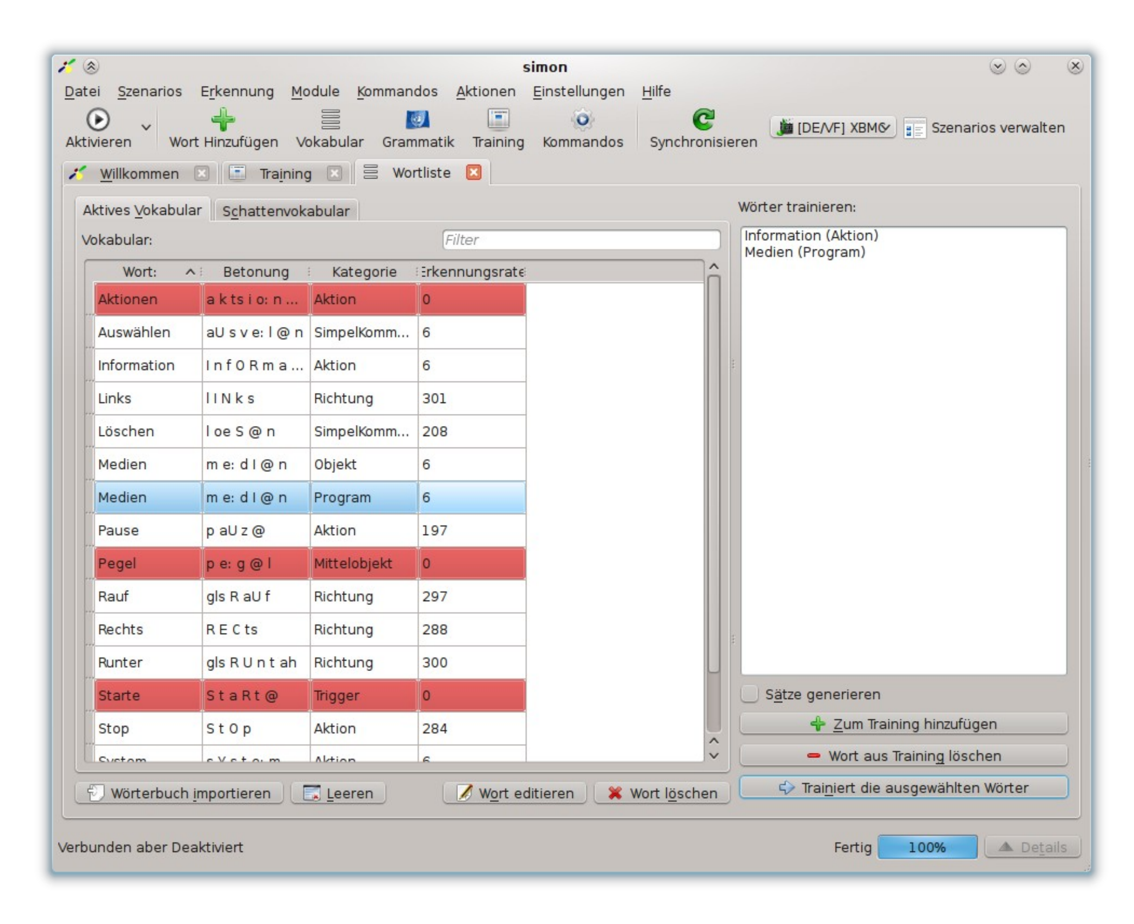


simon kann herkömmliche Textdateien importieren um Sie für das Training zu verwenden.

Spezialtraining

Zusätzlich zu den Trainingstexten erlaubt simon es auch individuelle Wörter oder Wortkombinationen aus Ihrem Lexikon direkt zu trainieren.

Diese Funktion befindet sich im Vokabel-Menü von simon.



Wählen Sie die Wörter zum Trainieren aus dem Vokabular links aus und ziehen Sie sie mit gedrückter Maustaste zur Liste rechts (Sie können das Wort auch in der Tabelle auswählen und dann auf "Zum Training hinzufügen" klicken).

Starten Sie das Training in dem Sie "Trainiert die ausgewählten Wörter" auswählen. Das Training selbst ist das gleiche wie bei gespeicherten Trainingstexten.



Wenn mehr als neun Wörter zu trainieren sind, wird simon automatisch den Text gleichmäßig auf mehr als eine "Seite" (Aufnahme) aufteilen.

Natürlich können Sie auch Wörter aus Ihrem Schattenlexikon zur Liste der zu trainierenden Wörter hinzufügen. Bevor das Training dann gestartet werden kann, müssen Sie aber auch hier alle Wörter zum aktiven Lexikon hinzufügen (siehe oben).

Trainingsdaten importieren

Mit der Trainingsdaten importieren Funktion können Sie zuvor gesammelte Trainingsdaten von früheren Versionen von simon importieren ohne das komplette Wörterbuch mitzuübernehmen.

Diese Funktion ist sehr speziell. Bitte benutzen Sie sie nur wenn Sie sich sicher sind, dass Sie wissen was Sie tun.

Sie können entweder seperate Promptsdateien angeben oder simon die Transkription aus den Dateinamen ablesen lassen.

Wenn Sie prompts basierte Transkriptionen verwenden (UTF-8 enkodiert) muss die Eingabedatei den folgenden inhalt besitzen: "[Dateiname] [Inhalt]". Dateinamen sind ohne Dateiwereiterung anzugeben und der Inhalt besteht nur aus Großbuchstaben. Zum Beispiel: Die Zeiele "demo_2007_03_20 DEMO" bedeutet dass die datei "demo_2007_03_20.wav" den Inhalt "Demo" besitzt.

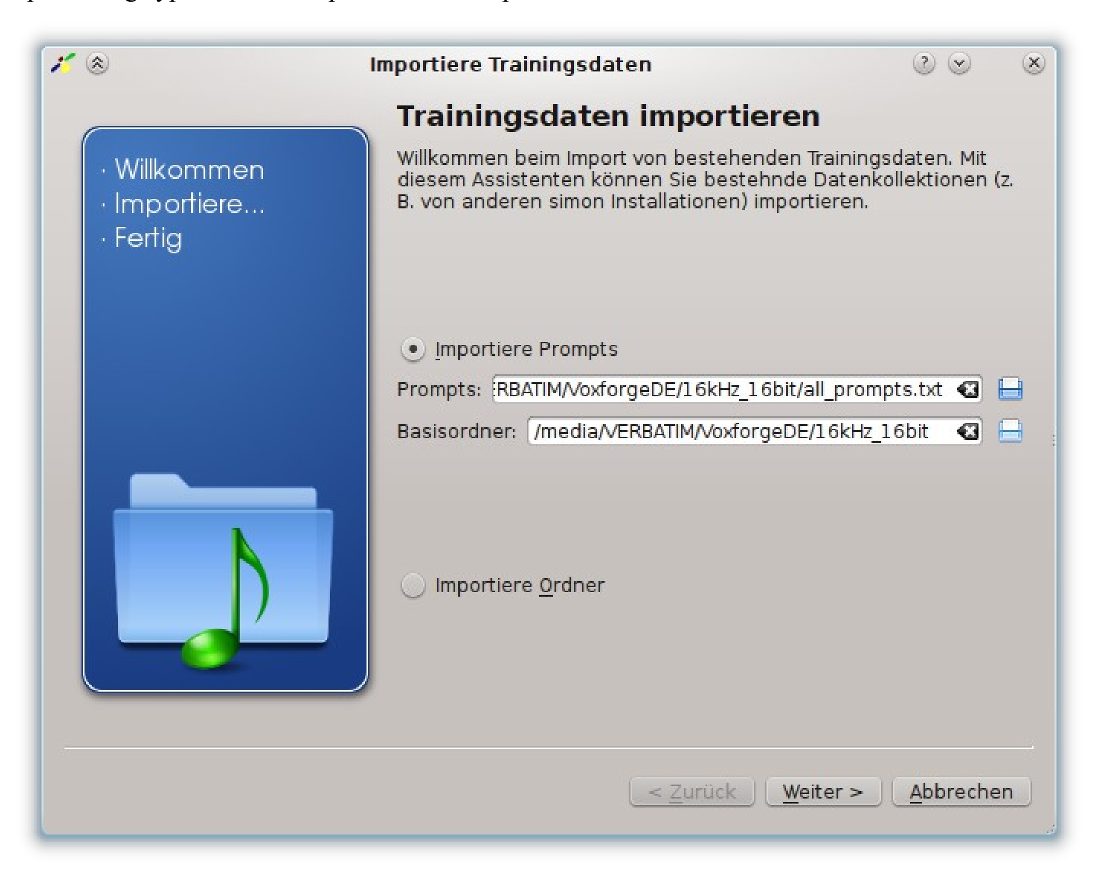
Weil Promptsdateien keine Dateierweiterung besitzen wird simon wav, mp3, ogg und flac probieren (in dieser Reihenfolge). Sollte eine Datei mit einem solchen Dateinamen existieren wird diese Datei importiert und folgende Übereinstimmung ignoriert (sollten demo.wav und demo.mp3 existieren wird nur die demo.wav importiert). Dies ist im Gegensatz zu Dateibasierten Transkriptionen wo beide Dateien importiert werden würden.

Wenn dateibasierte Transkriptionen verwendet werden, wird der Assistent die Transkription aus den gefunden Trainingsdaten anhand der *Dateinamen* auslesen. Eine Datei die `das_ist_ein_Text.wav` genannt ist, *muss* die Worte "das ist ein Test" beinhalten und nichts anderes. Satzzeichen wie (".", "-",...) in den Dateinamen werden ignoriert.

Dateien die mit simon 0.2 aufgenommen wurden, werden automatisch dieser Namensgebung folgen sodass Sie sie ohne Probleme importieren können sollten. Dateien Die von vorigen Versionen von simon aufgenommen wurden sollten nicht importiert werden. Sie können diese Aufnahmen aber natürlich über den Prompts-basierten Import importieren.

Die Dateien und Ihre Transkription werden dann zum Trainingskorpus hinzugefügt.

Um einen Ordner zu importieren der die Trainingsdateien beinhaltet wählen Sie ihn - und abhägig vom Importierungstyp auch die entsprechende Prompts-Datei - im Assistenten aus.



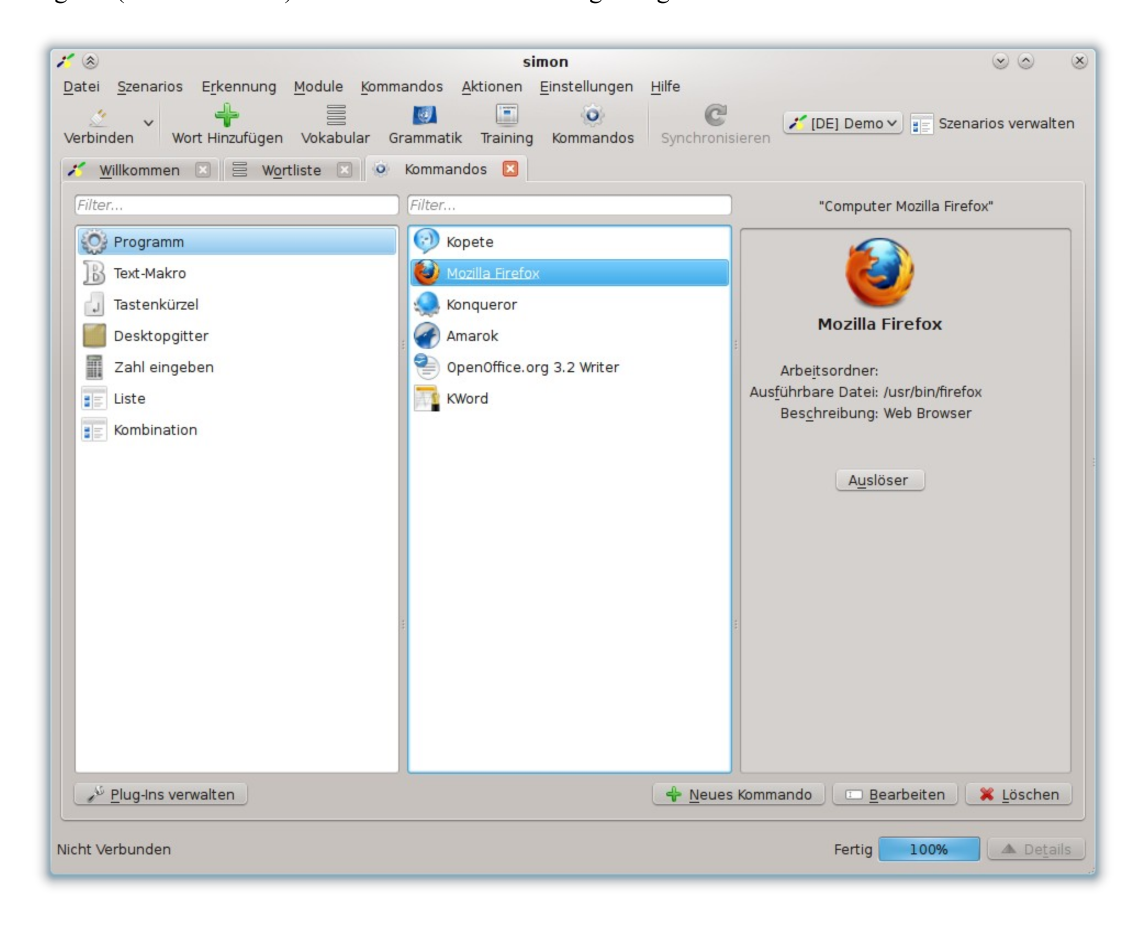
Der Ordner wird rekursiv durchsucht. Das heißt, dass der angegebene Ordner und all seine Unterordner nach `.wav`, `.flac`, `.mp3` und `.ogg` Dateien durchsucht wird. Alle gefundenen Dateien werden importiert.

Wenn Trainingsdaten importiert werden, werden alle konfigurierten Nachbearbeitungsfilter angewandt.

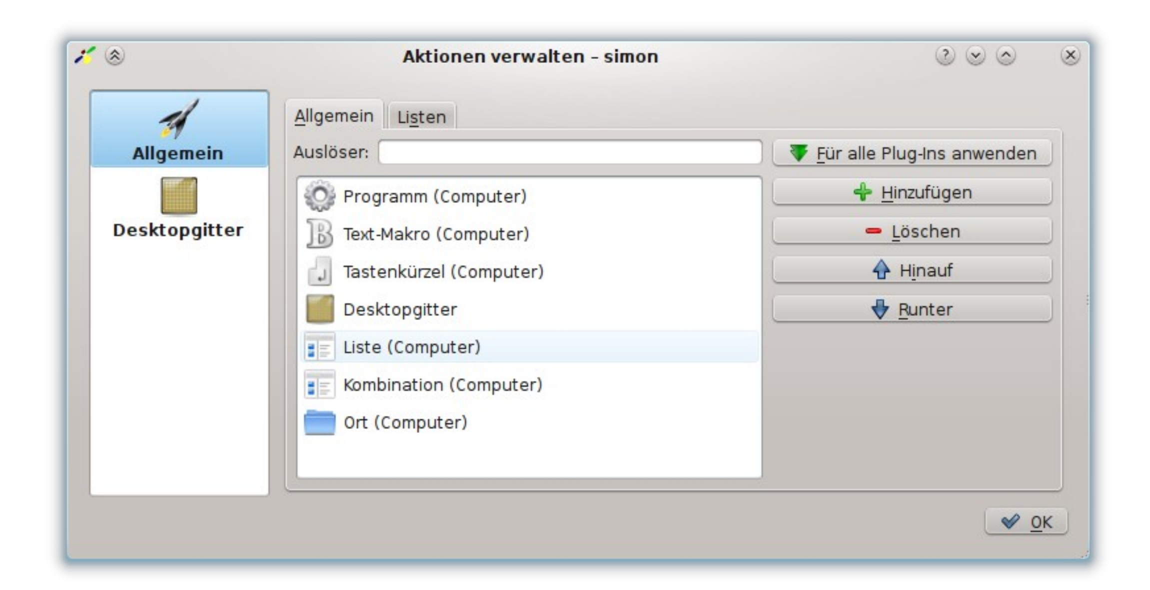
Wenn Sie etwas anderes als WAV Dateien importieren sind Sie dafür verantwortlich, dass die Dateien während dem Import korrekt dekodiert werden (zum Beispiel durch Nachbearbeitungsketten) oder die Modellerstellung *wird* fehlschlagen.

Kommandos

Wenn simon aktiviert ist und es erkennt etwas wird das Erkennungsergebnis den geladenen Kommando Plug-Ins (der Reihe nach) für die weitere Verarbeitung übergeben.



Das Kommandosystem kann mit einer Gruppe von Fabrikarbeitern verglichen werden. Jeder von Ihnen weiß, wie er genau eine Aufgabe bewältigt (z.B. Weiß "Karl" wie man Programme startet und "Hans" weiß wie Ordner geöffnet werden, etc.). Wenn simon jetzt etwas erkennt wird dieses Erkennungsergebnis zuerst an "Karl" übergeben, der sich das einmal ansieht. Ist die Anweisung für ihn gedacht, handelt er danach (und startet in diesem Beispiel ein Programm). Ansonsten gibt er die Anweisung weiter an "Hans", usw. Wenn keiner der geladenen Plug-Ins weiß was zu tun ist, wird das Erkennungsergebnis verworfen. Die Abfolge in der das Erkennungsergebnis zu den einzelnen Plug-Ins übergeben wird ist konfigurierbar in der Kommandokonfiguration (Kommandos > Plug-Ins verwalten).



Jedes Plug-In kann mit einem "Auslöser" verbunden werden. Mit Auslösern kann die Verantwortung der einzelnen Plug-Ins besser aufgeteilt werden.

Wenn wir wieder zu unseren Fabrikarbeitern zurückgehen könnten wir es vergleichen mit dem expliziten sagen des Namens der zustädigen Person. Statt "Öffne meine eigenen Dateien" sagen wir jetzt "Hans, öffne meine eigenen Dateien!" und "Hans" (das Plug-In das zustädig ist für das Öffnen der Ordner) weiß sofort, dass diese Anfrage für ihn ist.

In der Praxis bedeutet dass, z.B. dass wenn Sie ein Programm-Kommando "Firefox" haben um den Webbrowser zu öffnen und ein Ort-Kommando "Google" um die Websuche zu öffnen und die Auslöser "Starte" für das "Program" Plug-In und "Öffne" für das "Ort" Plug-In vergeben, müssten Sie "Starte Firefox" sagen um Firefox zu starten - statt einfach "Firefox" (den Namen des Kommandos) wenn Sie keinen Auslöser verwenden würden. Analog dazu müssten Sie "Öffne Google" sagen statt nur "Google" ohne Auslöser.

Trigger sind natürlich keine Voraussetzung und Sie können simon auch ohne Trigger verwenden (manche Kommandos haben aber per Default den Auslöser "Computer" den Sie in diesem Fall entfernen müssten). Aber sogar wenn Sie nur einen einzelnen Auslöser für all Ihre Plug-Ins verwenden (wie etwa den allgemeinen Trigger "Computer" um z.B. "Computer, Firefox" und "Computer Google" zu sagen) hat dies den großen Vorteil die versehentlichen Kommandoaufrufe zu minimieren.

simons Kommando Sektion zeigt den kompletten Satz der mit dem Kommando verbunden ist in der rechten oberen Ecke der Kommandoanzeige an.

Sie können mehrere Instanzen eines Plugins pro Szenario laden. Jede Instanz kann unterschiedliche Plugin Auslöser verwenden.

Jedes Kommando hat einen Namen (der das Kommando - mit einem etwaigen Plug-In Auslöser verbunden - auslöst), ein Symbol und mehrere Felder die vom Typ des Kommandos abhägen (siehe unten).

Einige Kommandoplugins bieten vielleicht eine Konfiguration für das Plugin selbst an (nicht die Kommandos die es verwaltet). Diese Konfigurationsseiten werden direkt in den Plug-Ins konfigurieren Dialog eingebunden (unter dem Allgemein Menüeintrag) wenn Sie das Plugin laden.

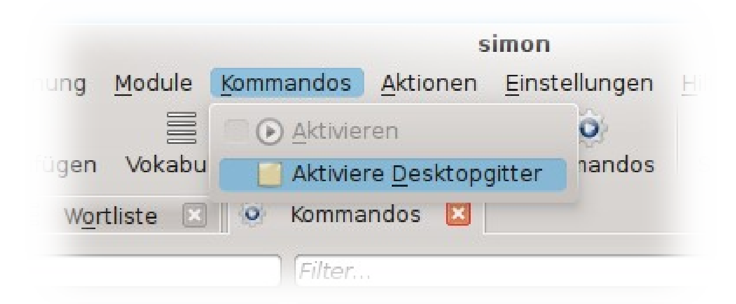
Plug-Ins die eine grafische Benutzeroberfläche anbieten (zum Beispiel das Nummer eingeben Kommandoplugin) können über "Sprachkommandos" konfiguriert werden. Sie können das assoziierte Wort ädern das den grafischen Knopf auslöst ädern aber auch das angezeigte Symbol, usw. Wenn Sie alle Sprachkommandos von einem grafischen Element entfernen wird das Element nicht mehr angezeigt (versteckt).

Sprachkommandos können gleich wie normale Kommandos durch die Kommandokonfiguration verwaltet werden.



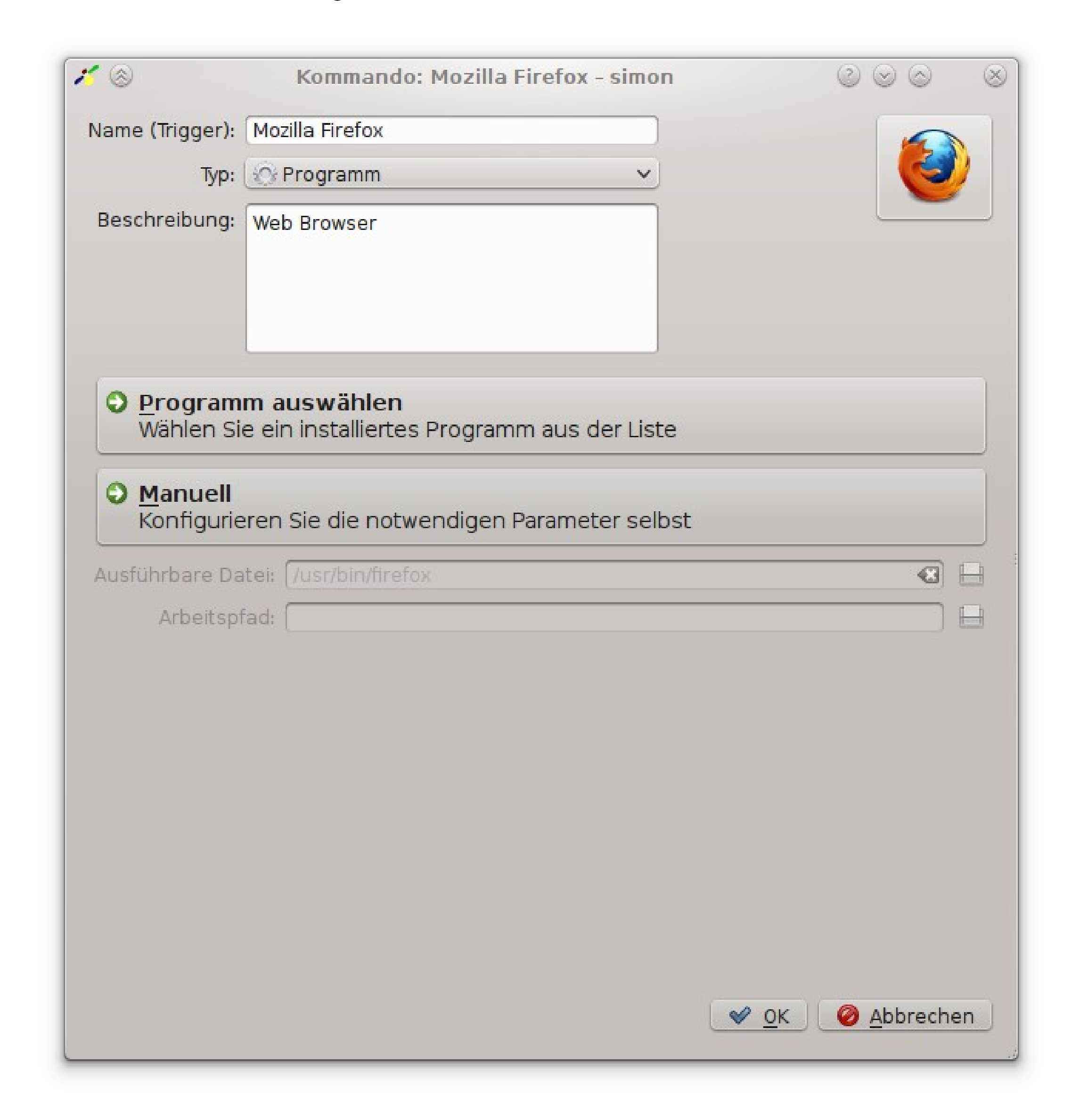
Um ein neues Sprachkommando zu einer Funktion hinzuzufügen, wählen Sie die Aktion die Sie mit dem Kommando verbinden wollen aus und klicken Sie Von Vorlage erzeugen. Passen Sie die erzeugten Werte an Ihre Wünsche an.

Einige Plug-Ins (zum Beispiel das Desktopgitter oder der Taschenrechner können auch einen Menüeintrag im Kommando Menü anzeigen.



Programm-Kommandos

Programm-Kommandos sind assoziiert mit einer ausführbaren Datei ("Programm") welches gestartet wird wenn das Kommando ausgelöst wird.



Argumente zu den ausführbaren Dateien werden unterstützt und an das Programm übergeben. Wenn der Pfad zur ausführbaren Datei oder die Argumente Leerzeichen beinhalten, *müssen* sie aber von Anführungszeichen umgeben werden.

Nehmen wir die ausführbare Datei `C:\Programme\Mozilla Firefox\firefox.exe" "C:\test datei.html"` an, wäre die korrekte Eingabe für das Feld "Ausführbare Datei" folgendes: `"C:\Programme\Mozilla Firefox\firefox.exe" "C:\test datei.html"`.

Der Arbeitsordner bestimmt in von welchem Ordner aus der Prozess gestartet wird. Wenn der Arbeitsordner `C:\folder` ist und das Kommando `"C:\Programme\Mozilla Firefox\firefox.exe" file.html` gestartet wrid, würde Firefox für die Datei

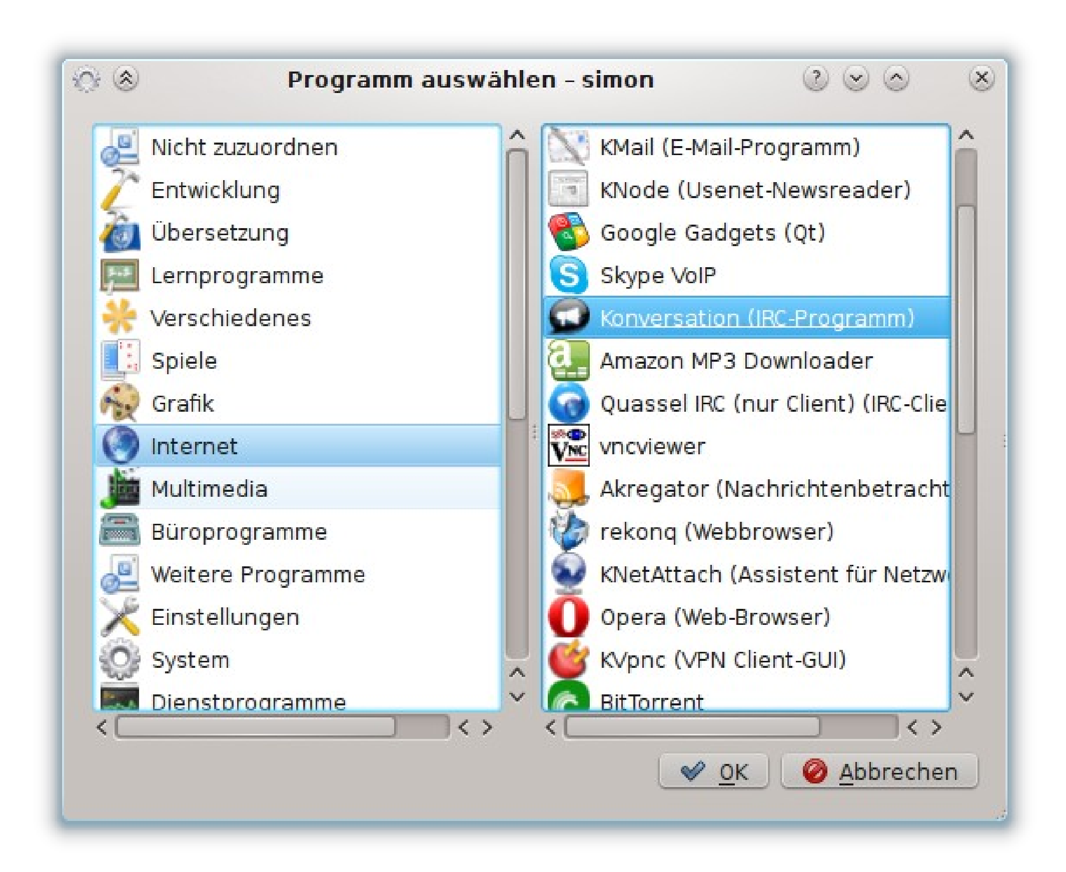
`C:\folder\file.html` suchen.

Der Arbeitsordner muss normalerweise nicht explizit gesetzt werden und kann leer gelassen werden.

Programme importieren

simon bietet einen Dialog mit dem Sie Programme auch direkt aus dem KDE Menü auswählen können.

> **Note:** Diese Option ist unter Microsoft Windows nicht verfügbar.



Der Dialog wird alle Programme aus dem KDE Menü in ihrer entsprechenden Kategorie anzeigen.

Unterkategorien werden nicht unterstützt und werden auf der selben Ebene wie ihre übergeordneten Kategorien angezeigt.

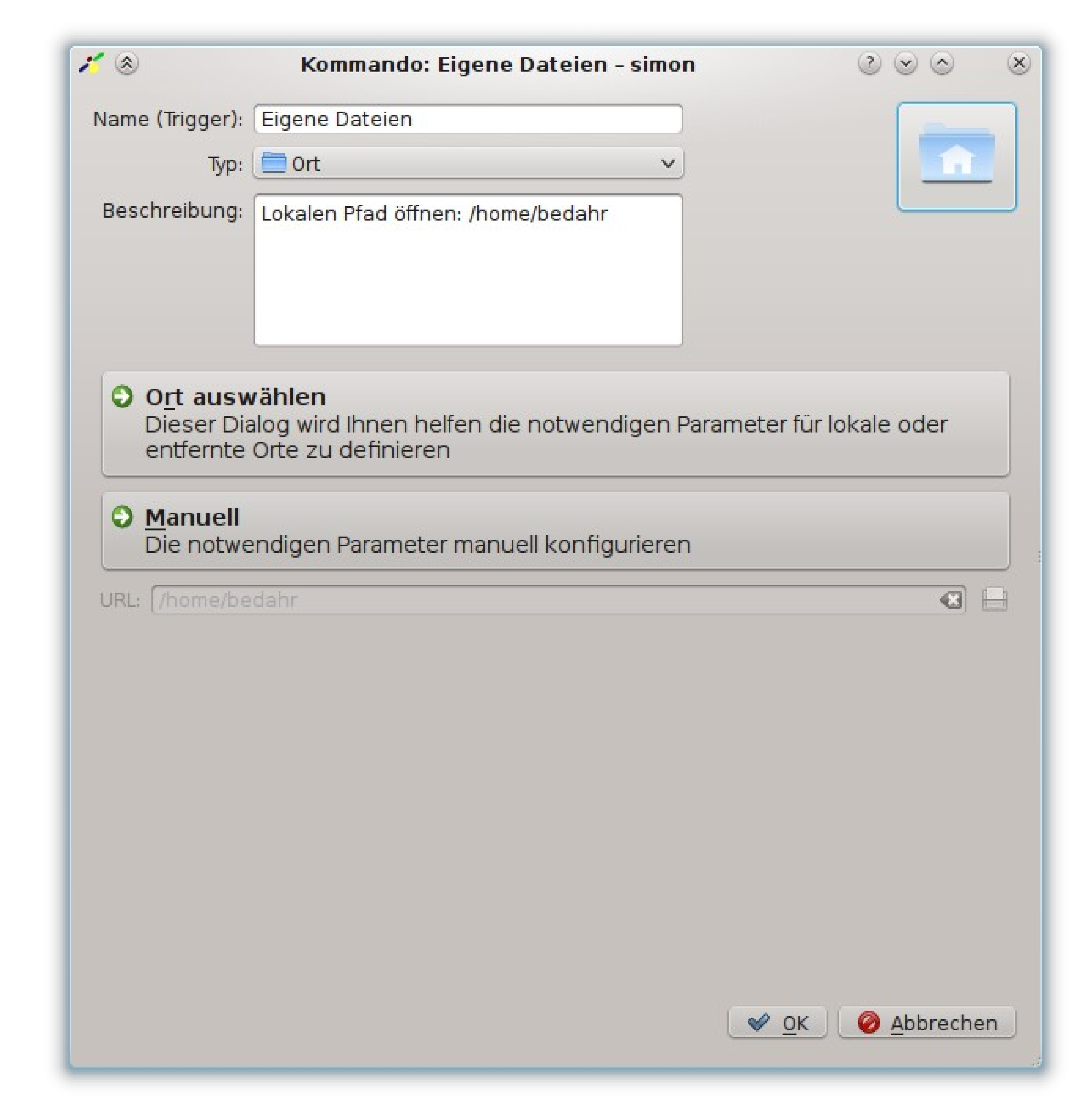
Wählen Sie einfach das Programm aus, das Sie importieren wollen und schließen klicken Sie Ok. Die korrekten Werte für die Ausführbare Datei und für den Arbeitsordner werden automatisch ausgefüllt.

Ort-Kommandos

Mit Ort Kommandos kann simon einen beliebigen URL öffnen. Weil simon die Adresse einfach an den zugrunde liegenden URL-Handler der jeweiligen Plattform weitergibt werden alle von diesem Unterstützten Protokolle wie "remote:/" (Linux/KDE spezifisch) oder sogar KDEs "Web-Shortcuts" unterstützt.

Neben herkömmlichen Ordnern können aber auch Dateien angegeben werden die dann mit dem mit

ihnen standardmäßig assoziierten Programm geöffnet werden wenn das Kommando ausgelöst wird.

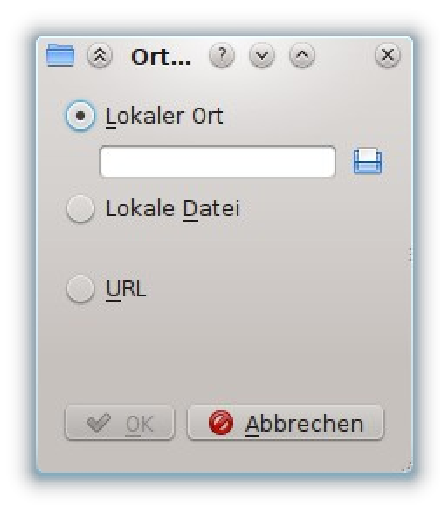


Um einen URL mit einem Kommando zu verbinden können Sie den URL entweder manuell in das URL-Feld eingeben (wählen Sie Manuell) oder ihn mit dem Ort importieren Dialog auswählen.

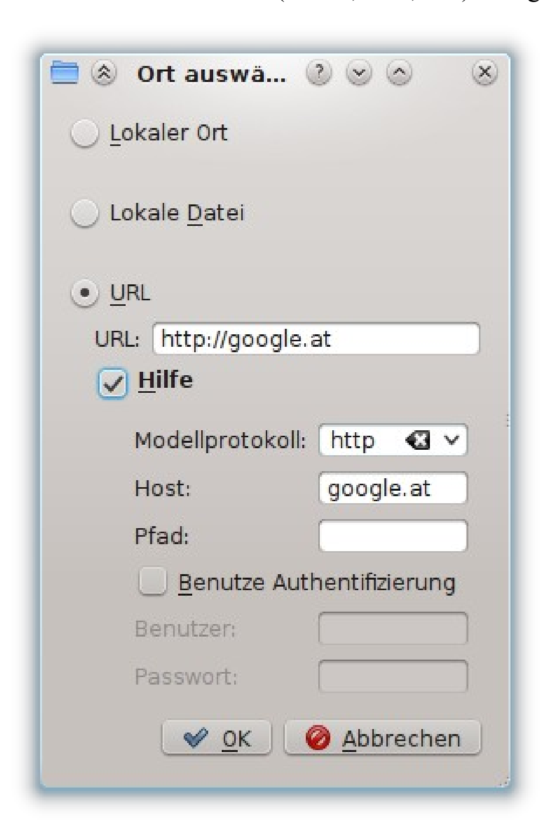
Orte importieren

Der Ort importieren Dialog erleichtert es eine korrekte URL anzugeben.

Um einen lokalen Ordner oder eine lokale Datei anzugeben, wählen Sie Lokaler Ort und wählen Sie den Ordner oder die Datei mit dem Feld aus.



Um einen entfernten URL (HTTP, FTP, etc.) anzugeben wählen Sie URL.

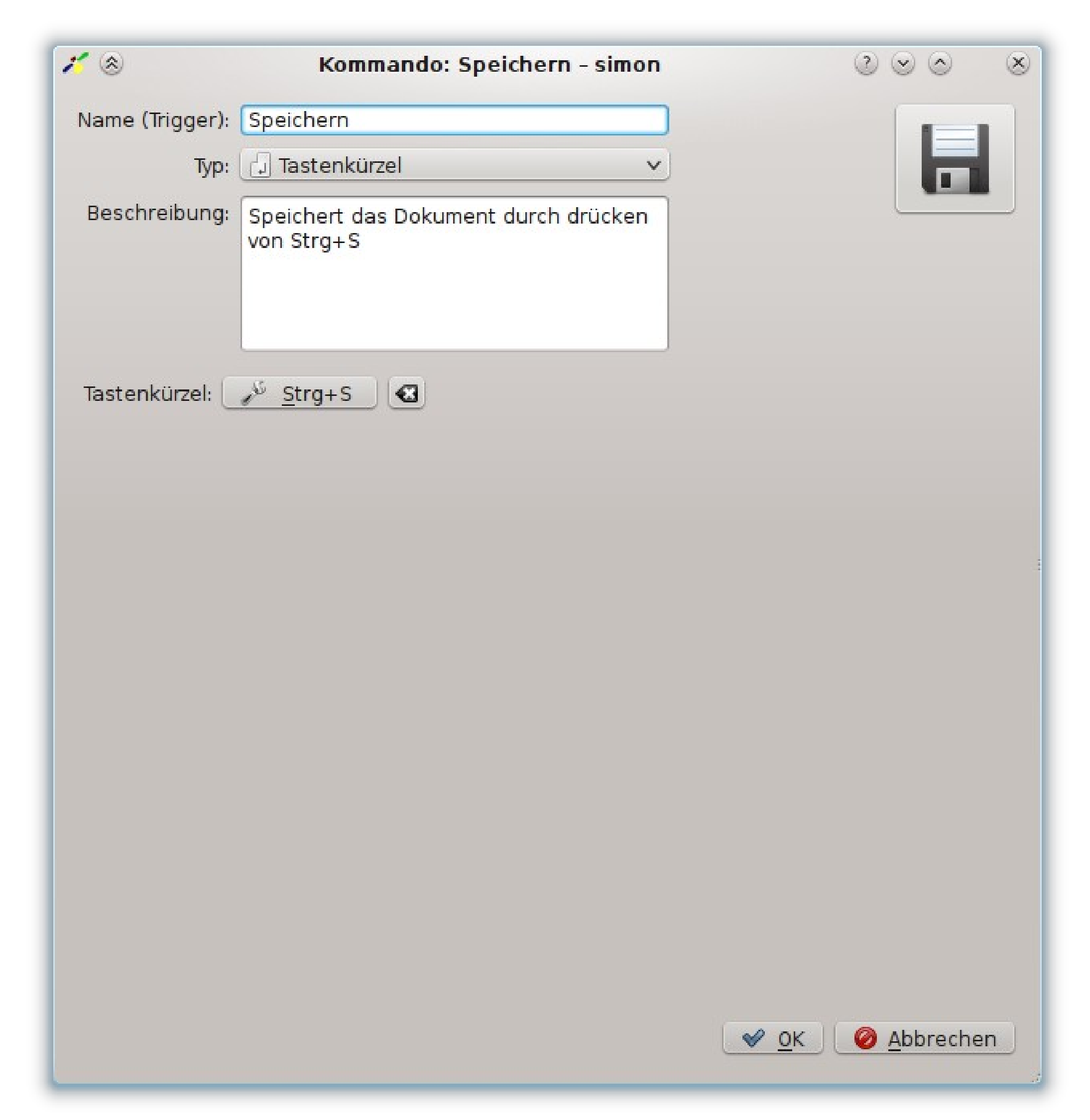


Bitte beachten Sie, dass URLs mit Authentifizierungsinformation im Klartext gespeichert werden und das Passwort also unverschlüsselt vorliegt.

Tastenkürzel-Kommandos

Mit den Tastenkürzel-Kommandos kann ein Benutzer Kommandos mit Tastenkombinationen verbinden.

Das Kommando kann Tasteneingaben simulieren um Tastenkürzel wie "Strg+C"oder "Alt-F4" auszulösen.



Um ein Tastenkürzel auszuwählen, klicken Sie einfach auf den Tastenkürzel-Knopf und drücken anschließend eben jene Kombination die Sie mit dem Kommando assozieren möchten auf Ihrer Tastatur.

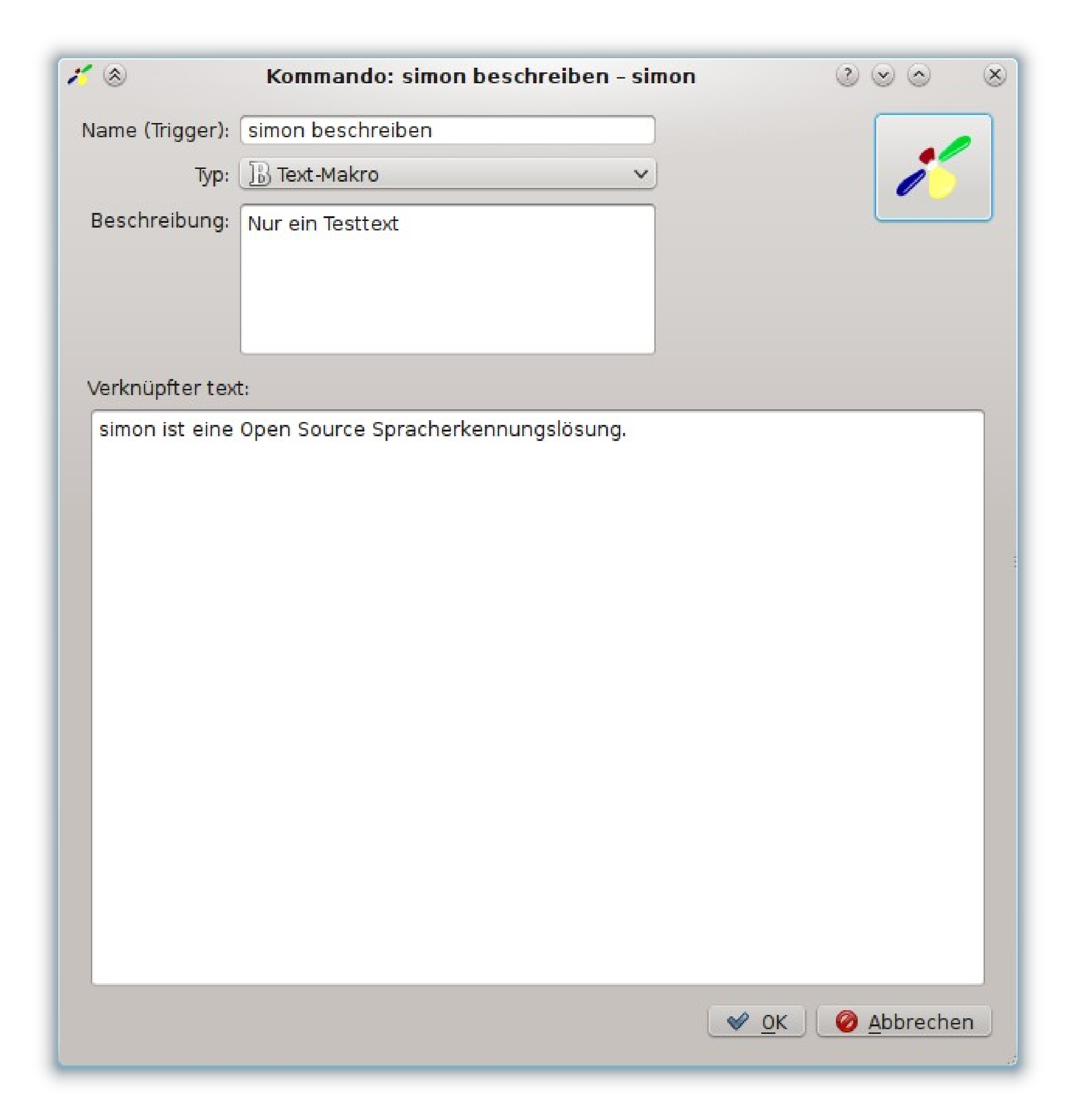
simon wird das Tastenkürzel abfangen und mit dem Kommando assozieren.

Wegen technischen Einschräkungen gibt es unter Microsoft Windows einige Tastenkürzel die sich nicht ohne weiteres abfangen lassen mit simon (z.B. Strg+Alt+Entf und Alt+F4). Diese speziellen Tastenkombinationen können mit einer eigenen Liste ausgewählt werden.

> **Note:** Diese Auswahlliste ist im Bildschirmfoto nicht sichtbar da sie nur unter Microsoft Windows angezeigt wird.

Text-Makro-Kommandos

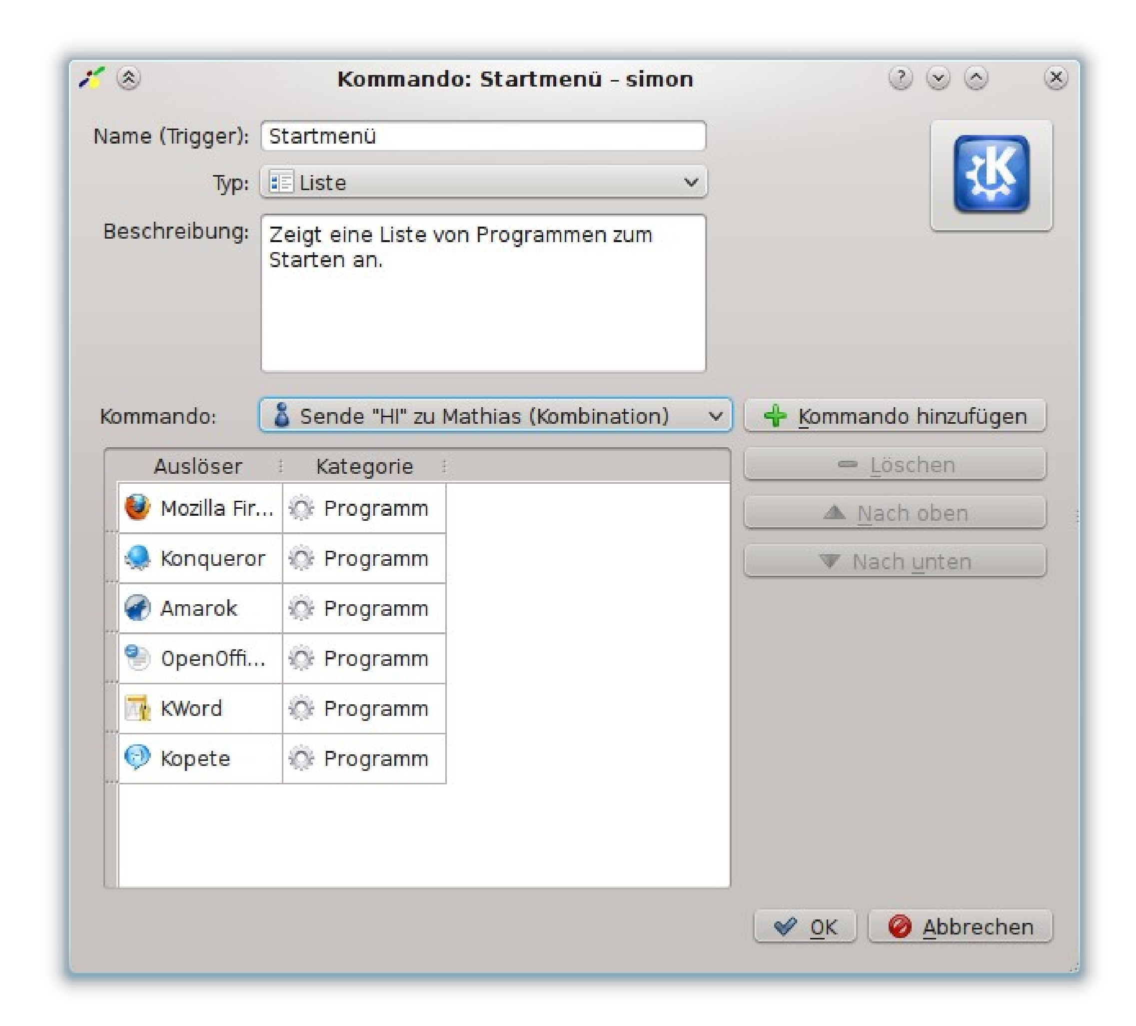
Mit Text-Makro-Kommandos kann der Benutzer Text mit Kommandos assozieren. Wenn ein Text-Makro-Kommando ausgelöst wird, wird der Text mit dem das Kommando verbunden ist über simulierte Tastendrücke "geschrieben".



List-Kommandos

List-Kommandos kombinieren mehrere Kommandos (alle Typen von Kommandos werden unterstützt) in eine einzelne Liste. Der Benutzer kann dann einen Eintrag der Liste auswählen indem er dessen Nummer (1-9) sagt und damit dieses Kommando auslösen.

List-Kommandos sind sehr hilfreich um die Menge an Wörter im Vokabular und damit die Menge der benötigten Trainingsdaten gering zu halten.

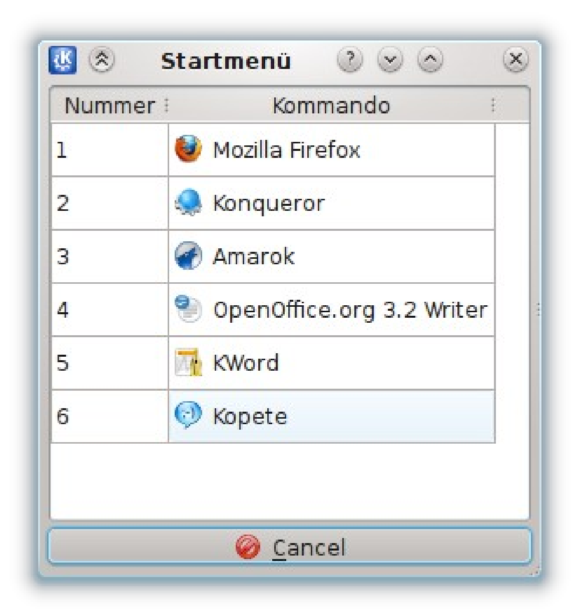


List-Kommandos sind vor allem dann nützlich wenn Kommandos schwierig auszusprechende Trigger haben oder mehrere Kommandos logisch gruppiert werden können. Ein typisches Beispiel wäre ein Kommando "Startmenü" das eine Liste von Programmen (Program-Kommandos) auflistet. So können diese Program-Kommandos sehr komplexe aber beschreibende Namen besitzen (wie z.B.: "OpenOffice.org Writer 3.1") ohne dass simon diese Wörter jemals erkennen muss um das Kommando auszulösen.

Kommandos unterschiedlicher Typen können in einer Liste gemischt werden.

Listen Kommandos Anzeige

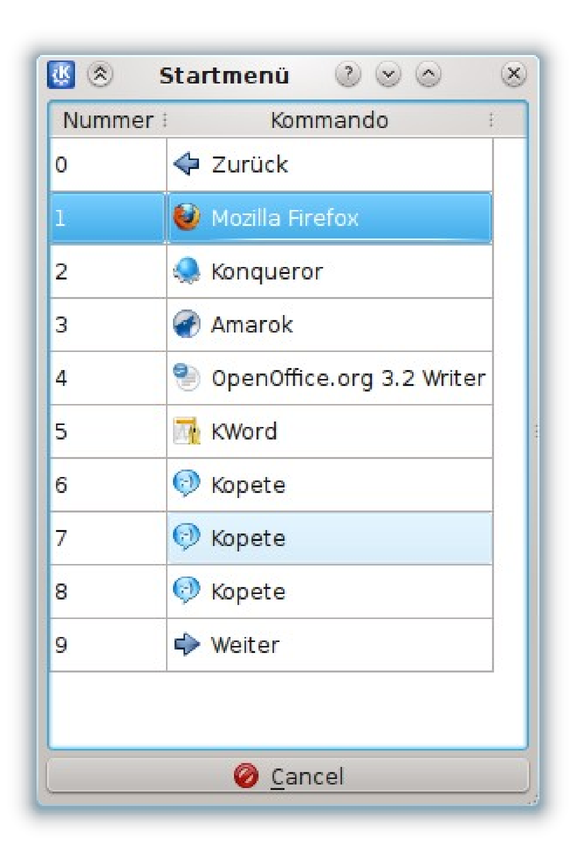
Wenn ein List-Kommando ausgelöst wird, wird eine Liste in der Mitte des Bildschirmes angezeigt. Diese Liste wird automatisch größer und kleiner um alle Elemente anzeigen zu können.



Der Benutzer kann dann die Kommandos auslösen in dem er die Nummer sagt die mit dem Kommando verbunden ist, dass er auslösen möchte (In diesem Beispiel: "Eins" um Mozilla Firefox aufzurufen).

Während eine Liste aktiv ist (sie wird angezeigt) wird jedes Erkennungsergebnis das nicht die Liste selbst betrifft (andere Kommandos, etc.) verworfen. Der Auswahlprozess kann jederzeit abgebrochen werden durch sagen von "Abbrechen"oder durch klicken auf das "Abbrechen" Feld.

Wenn mehr als 9 Kommandos mit der Liste assoziiert sind, werden automatisch "Weiter" und "Zurück" Elemente angezeigt die ebenfalls mit Nummern versehen sind ("Null" wird mit "Zurück" und "Neun" wird mit "Weiter" verbunden).



Konfiguration von Listenelementen

Standardmäßig verwendet das Listenkommando die folgenden Auslöser.Um List-Kommandos zu ihrem vollen Potential zu benutzen, stellen Sie sicher, dass Ihr Sprachmodell die folgenden "Sätze" kennt und erlaubt:

- "Null"
- "Eins"
- "Zwei"
- "Drei"
- "Vier"
- "Fünf"
- "Sechs"
- "Sieben"
- "Acht"
- "Neun"
- "Abbrechen"

Natürlich können Sie diese Wörter in der simon Konfiguration anpassen:

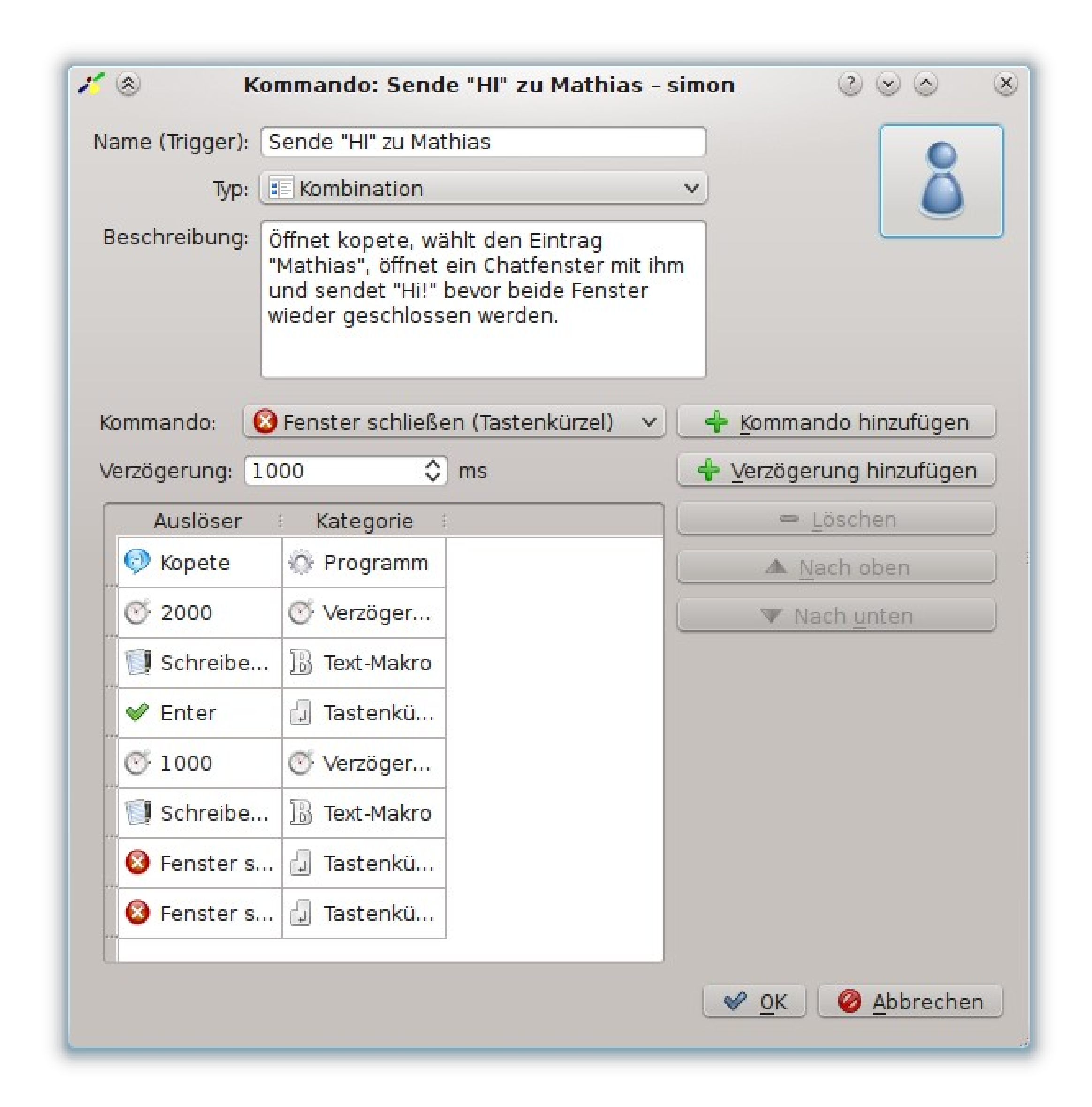
- Kommandos > Plug-Ins verwalten > Allgemein > Listen für die szenarioweite Listenkonfiguration.
- Einstellungen > simon einrichten... > Aktionen > Listen fürdie globale Konfiguration. When Sie ein neues Szenario anlegen wird diese Konfiguration als die Listenkonfiguration des neuen Szenarios übernommen.

Listenkommandos sind intern auch verwendet von anderen Plugins wie z.B. das Desktopgitter. Die Konfiguration der Trigger beeinflusst auch deren Listen.

Kombination-Kommandos

Kombination-Kommandos erlauben dem Benutzer mehrere Kommandos automatisch hintereinander ablaufen lassen.

Wenn ein Kombinationskommando ausgelöst werden, werden alle mit ihm assoziierten Kommandos in einer Reihenfolge die der Benutzer festlegt ausgeführt. Es können auch Wartezeiten zwischen Kommando-Aufrufen eingefügt werden.



Mit Kombination-Kommandos kann der Benutzer komplexe Arbeitsabläufe in ein einzelnes Kommando packen. Zum Beispiel:

- Starte Kopete (Program Kommando)
- Warte 2000ms bis Kopete gestartet ist
- Schreibe "Mathias" (Text-Makro-Kommando); Dies wird den Kontakt "Mathias"in meiner Kontaktliste auswählen
- Drücke Eingabe (Tastenkürzel-Kommando)
- Warte 1000ms bis Kopete das Chatfenster öffnet
- Schreibe "Hallo" (Text-Makro-Kommando) und drücke Eingabe (Tastenkürzel-Kommando) um die Nachricht zu senden.
- Drücke Alt+F4 (Tastenkürzel-Kommando) um das Chatfenster zu schließen.

- Drücke Alt+F4 (Tastenkürzel-Kommando) um das Kopete Hauptfenster zu schließen.

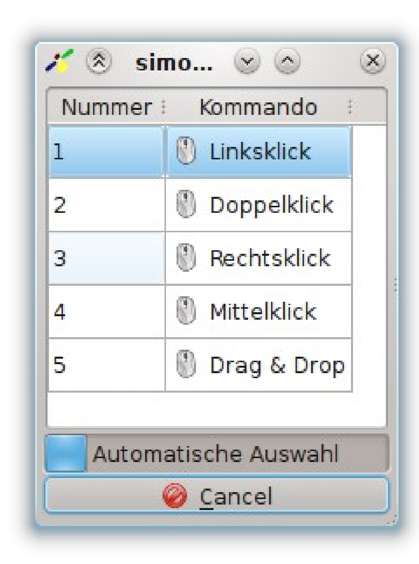
Desktopgitter

Das Desktopgitter erlaubt dem Benutzer die Maus mit der Sprache zu steuern.



Das Desktopgitter teilt den Bildschirm in neun Teile die von 1-9 durchnummeriert werden. Wenn eine der neun Nummer ausgesprochen wird, wird das Feld das mit dieser Nummer verbunden war in neun weitere Felder zerteilt die wieder von 1-9 durchnummeriert werden, etc. Dies wird insgesamt drei Mal gemacht. Nach der vierten Auswahl einer Zahl (der immer kleiner werdenden Teile des Bildschirms) wird in die Mitte des verbleibenden Vierecks geklickt (mit der linken Maustaste).

Die Exakte Aktion die beim "Klick" ausgewählt wird ist konfigurierbar. Standardmäßig wird der benutzer jedes mal durch eine Liste gefragt. Wird Drag and Drop ausgewählt wird das Desktopgitter nocheinmal angezeigt um den Punkt des "loslassens" auszuwählen.



Während das Desktopgitter aktiv ist (es wird angezeigt) wird jedes Erkennungsergebnis das nicht das

Desktopgitter selbst betrifft (andere Kommandos, etc.) verworfen. Der Auswahlprozess kann jederzeit abgebrochen werden durch sagen von "Abbrechen".

Das Plug-In "Desktopgitter" registriert eine Konfigurationsseite in der Kommandokonfiguration.



Der Auslöser, der das Desktopgitter aktiviert (anzeigt) ist natürlich frei konfigurierbar. Außerdem kann der Benutzer zwischen "echter" und "falscher" Transparenz unterscheiden. Wenn Ihre grafische Oberfläche Compositing Effekte erlaubt, kann simon "echte" Transparenz verwenden. Sollte echte Transparenz nicht möglich sein, wird simon mit "falscher" Transparenz arbeiten: Dabei schießt simon eine Momentaufnahme des aktuellen Bildschirmes bevor das Desktopgitter angezeigt wird und setzt dieses Bild als Hintergrundbild des Desktopgitters. Wenn verfügbar, sollte immer die "echte" Transparenz verwendet werden.

Wenn das Desktopgitter konfiguriert wurde echte Transparenz zu verwenden und das System unterstützt keine Compositing Effekte wird der Hintergrund des Gitters undurchsichtig und grau sein.

Beinahe alle aktuellen Systeme werden aber Compositing per Default unterstützen.

Das beinhaltet zum Beispiel:

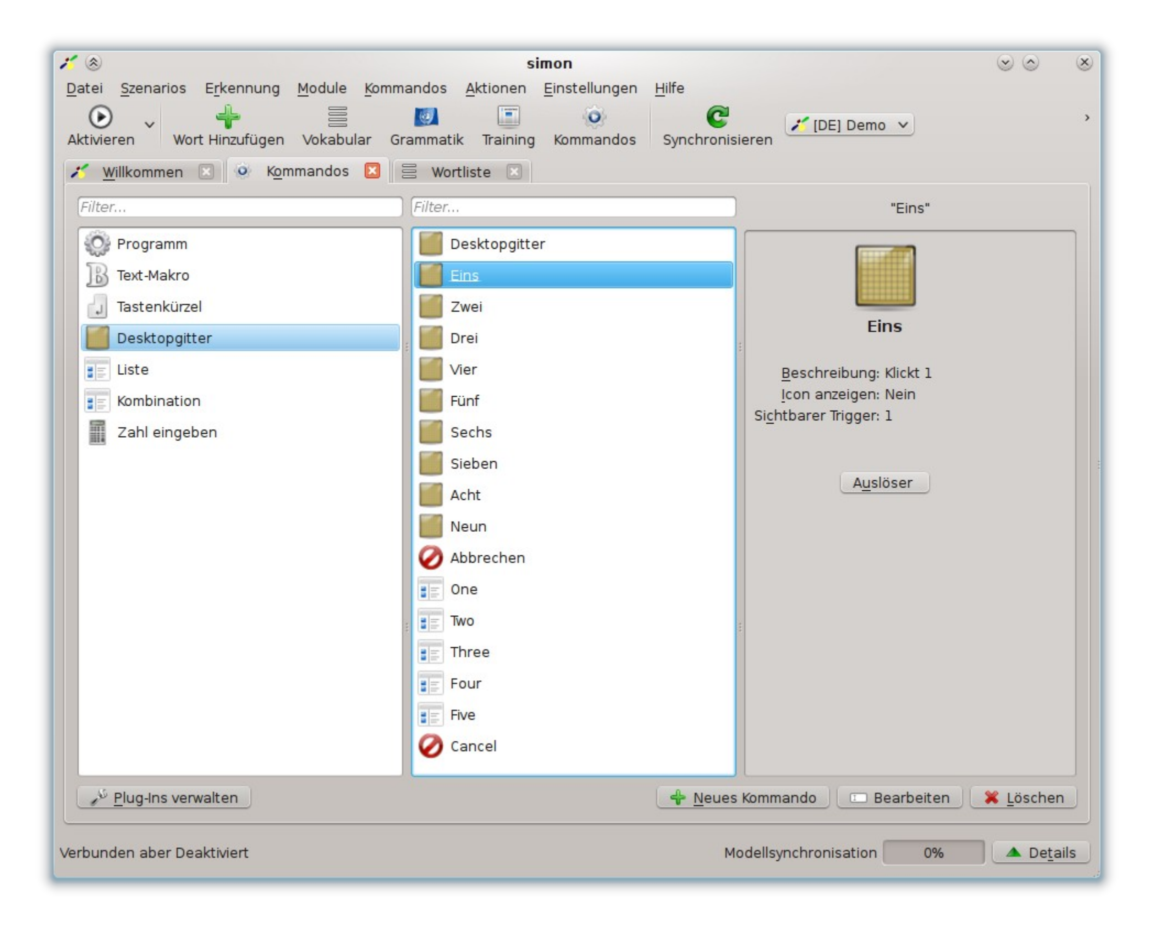
- Microsoft Windows 2000 oder höher (XP, Vista, 7)
- GNU/Linux mit einem Composite Manager wie Compiz, KWin4, xcompgrc, etc.

Standardmäßig verwendet das Desktopgitter nummern um die individuellen Felder auszuwählen.Um das Desktopgitter zu benutzen stellen Sie sicher, dass Ihr Sprachmodell folgende "Sätze" ermöglicht:

- "Eins"
- "Zwei"
- "Drei"
- "Vier"
- "Fünf"
- "Sechs"

- "Sieben"
- "Acht"
- "Neun"
- "Abbrechen"

Um diese Auslöser zu konfigurieren, verädern Sie einfach die Kommandos die mit dem Plugin assoziiert sind.



Nummer eingeben

Mit dem Nummer eingeben Plug-In kann der Benutzer leicht Nummern beliebiger Größe eingeben.

Mit dem Diktations- oder Text-Makro-Plug-In könnte man eigentlich leicht Nummern Ihren Ziffern zuweisen und das als Eingabemethode verwenden. Um höhere Nummern einzugeben gibt es dann zwei Möglichkeiten - jeweils mit entsprechenden Nachteilen:

- *Die vollen Wörter hinzufügen: "Elf", "Zwölf", etc.*

  Obwohl das als die eleganteste Lösung erscheint (weil sie es ermöglicht z.B.: "fünfhundertzweiundsiebzig" zu sagen) können wir leicht sehen das es ein enormer Aufwand wäre all diese Wörter hinzuzufügen - und sie dann auch noch zu trainieren. Und wo hören wir auf?

- *Die Nummern mit Ihren einzelnen Ziffern buchstabieren*

Obwohl das nicht ganz so elegant scheint wie die erste Lösung ist das natürlich viel praktikabler.

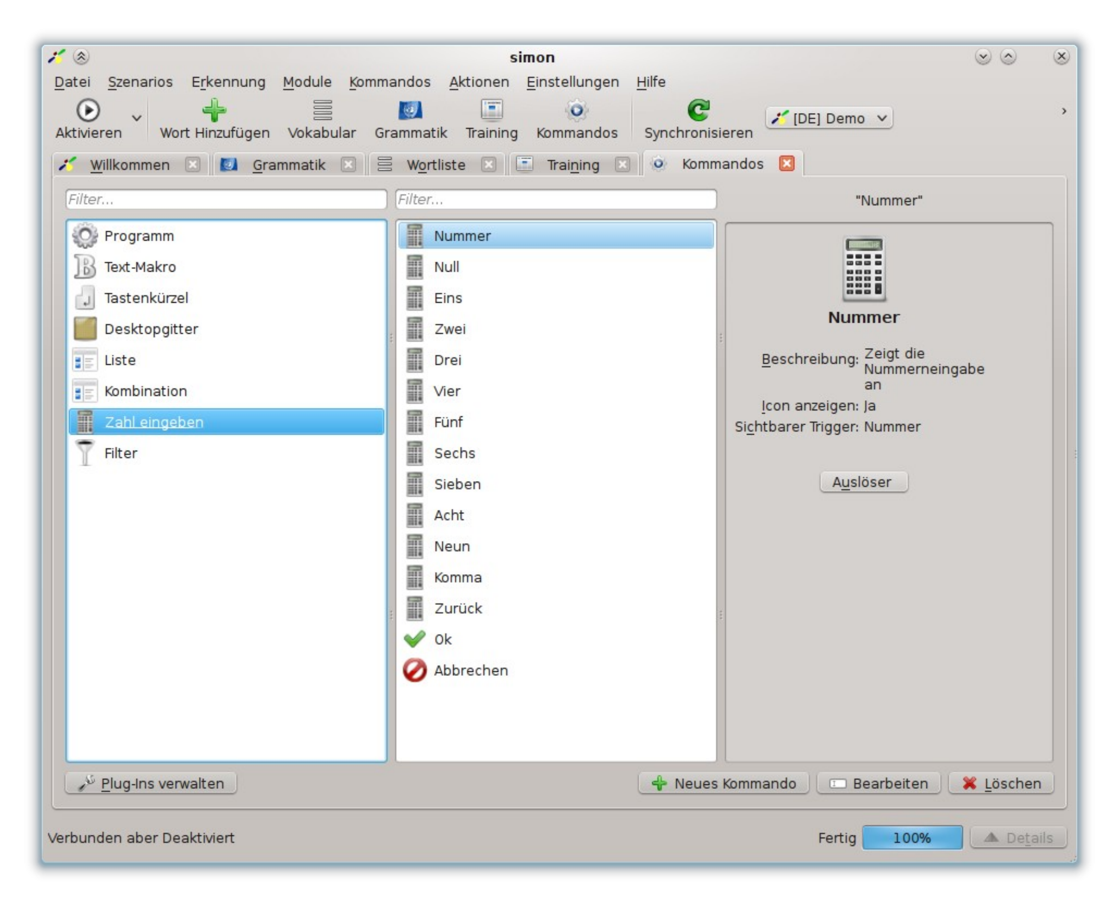
Leider erwarten einige Applikationen (wie z.B. das hervorragende mousless browsing firefox addon), dass der Benutzer eine Zahl zwar über die Tastatur ziffernweise eintippt aber dabei nur kurze Pausen zwischen den einzelnen Ziffern macht (das mouseless browsing Addon bspw. wird per Default maximal 500ms also eine halbe Sekunde zwischen den einzelnen Tastendrücken warten. Wenn Sie also bspw. 52 eingeben wollten würden Sie folgendes sagen: "Fünf (Pause) Zwei". Weil die Pause benötigt wird um den Satz zu beenden würde eine zeitgesteuerte Anwendung die Eingabe von "Fünf" als bereits vollstädig interpretieren.

Das Nummer eingeben Plug-In - wenn ausgelöst - zeigt eine Taschenrechner ähnliche Oberfläche um eine Zahl einzugeben. Die Eingabe kann jederzeit korrigiert werden indem der Benutzer "Zurück" sagt. Ein Dezimal-Komma kann eingefügt werden indem der Benutzer "Komma" sagt. Wenn "Ok" gesagt wird, wird die Zahl eingetippt. Da die tatsächliche Eingabe innerhalb des Plug-Ins stattfindet erhält die Applikation schlussendlich nur noch eine Serie von schnell hintereinander simulierten Tastendrücken die nur noch Millisekunden auseinanderliegen.



Während das Nummer eingeben Plug-In aktiv ist (es wird angezeigt) wird jedes Erkennungsergebnis das nicht da Nummer eingeben Plug-In selbst betrifft (andere Kommandos, etc.) verworfen. Die Eingabe kann jederzeit abgebrochen werden durch sagen von "Abbrechen"oder durch klicken auf das "Abbrechen" Feld.

Weil keine Kommandos vom Typ "Nummer eingeben" angelegt werden können, ist dieses Plug-In auch nicht im "Neues Kommando" Dialog aufgeführt. Das Nummer eingeben Plug-In registriert aber eine Konfigurationsseite in der Kommandokonfiguration.



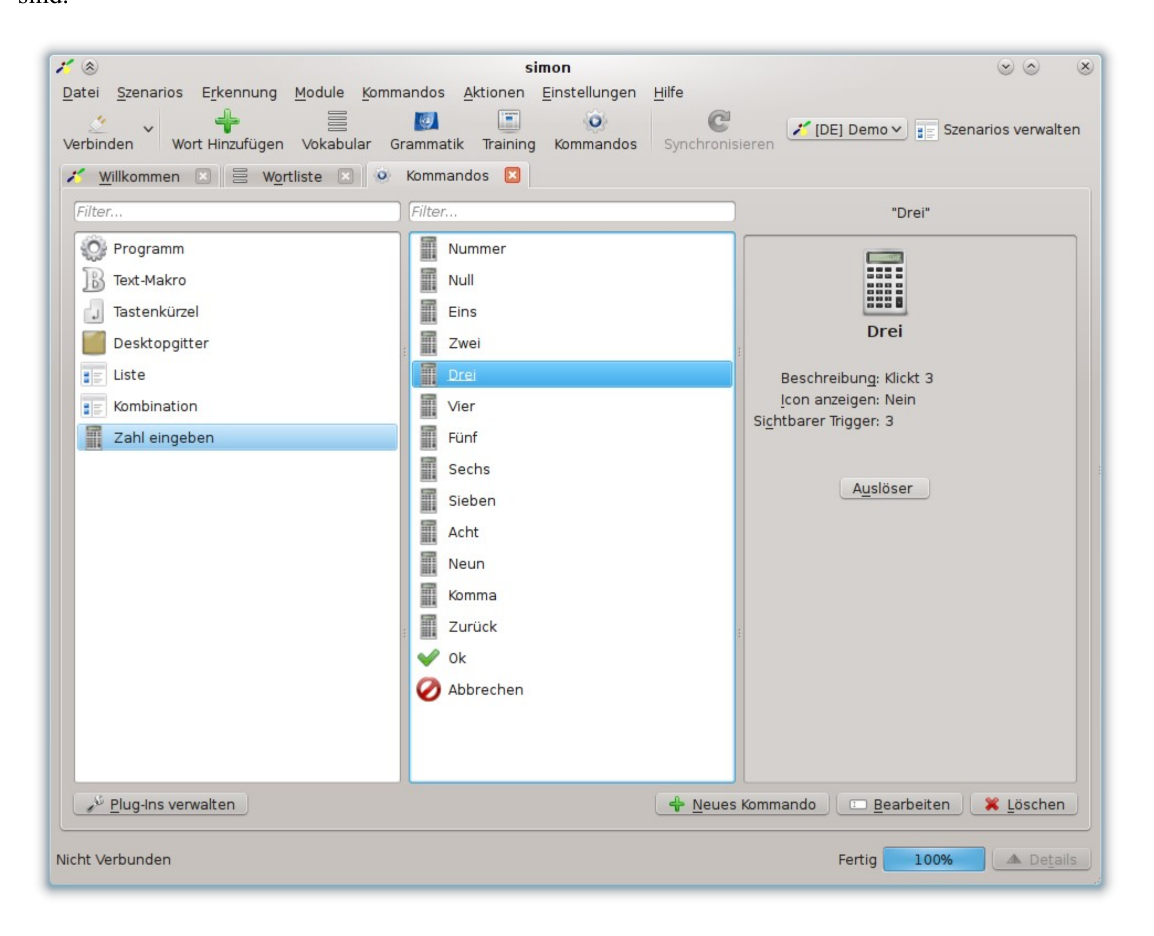
Der Auslöser definiert welches Wort oder welche Phrase die Oberfläche zum Eingeben der Nummer anzeigen wird.

Standardmäßig verwendet das Nummer eingeben Plug-In Nummern um die einzelnen Zahlen auszuwählen und einige Kontrollworte.Um das Nummer eingeben Plug-In komplett zu verwenden, stellen Sie bitte sicher, dass Ihr Sprachmodell folgende Sätze ermöglicht:

- "Null"
- "Eins"
- "Zwei"
- "Drei"
- "Vier"
- "Fünf"
- "Sechs"
- "Sieben"
- "Acht"
- "Neun"
- "Zurück"
- "Komma"

- "Ok"
- "Abbrechen"

Um diese Auslöser zu konfigurieren, verädern Sie einfach die Kommandos die mit dem Plugin assoziiert sind.



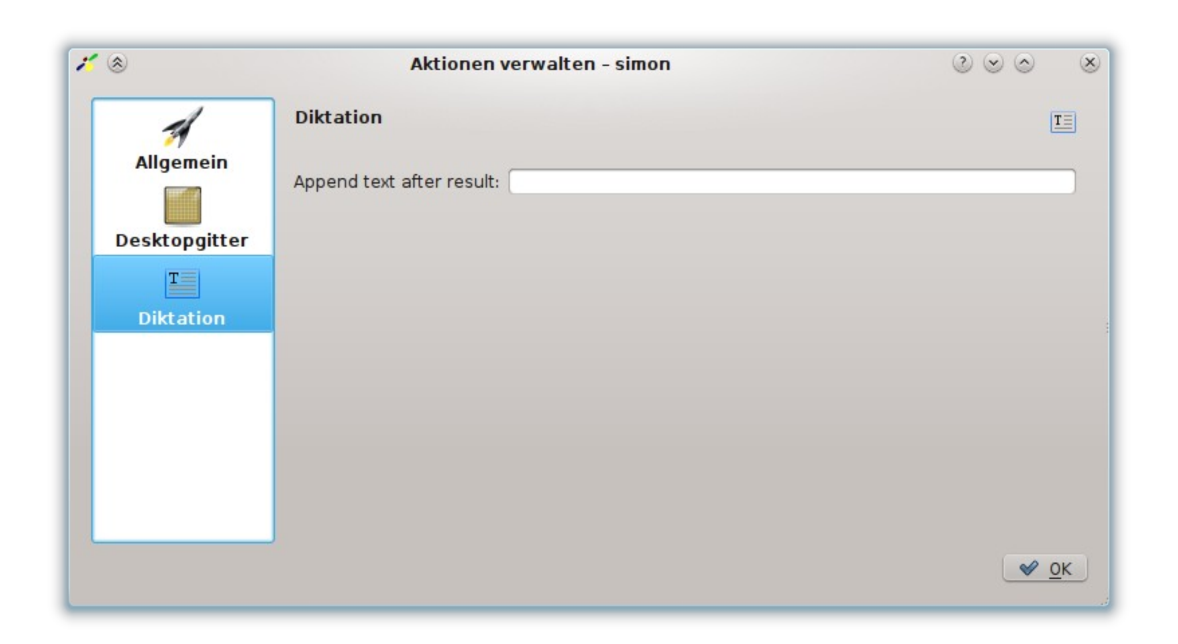
Diktation

Das Diktationsplug-in schreibt das Erkennungsergebnis mit simulierten Tastendrücken hin.

Wenn das Diktationsplug-in ohne Auslöser verwendet wird, wird das Plug-In alle Erkennungsergebnisse akzeptieren die es bekommt und diese einfach hinschreiben. Das heißt, dass alle Plug-Ins die *unter* einem Diktationsplug-in ohne Auslöser gereiht sind werden dadurch nie ein Erkennungsergebnis erhalten.

Weil keine Kommandos vom Typ "Diktation" angelegt werden können, scheint das Diktationsplug-in nicht im "Neues Kommando" Dialog auf.

Das Diktationsplugin kann so konfiguriert werden beliebigen Text nach jedem Erkennungsergebnis anzuhägen.

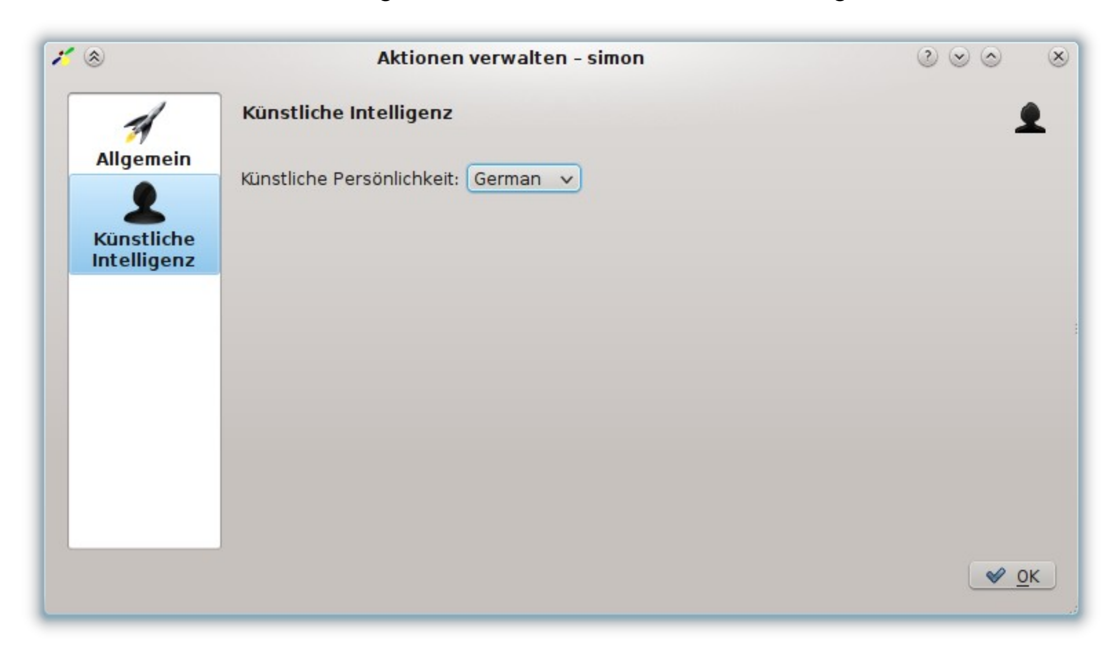


Künstliche Intelligenz

Das künstliche Intelligenz Plug-In ist ein Plug-In, dass eine menschliche Kommunikation simuliert.

Die KI (künstliche Intelligenz) verwendet das festival text to speech um Sprache auszugeben und somit mit dem Benutzer zu "sprechen" (Fragen beantworten oder z.B. über das Wetter tratschen).

Das Plug-In benutzt AIMLs für die "Intelligenz". Der Großteil der AIML Sets sollten unterstützt sein. Der bekannte A. L. I. C. E. Bot (http://www.pandorabots.com/pandora/talk?botid=f5d922d97e345aa1) und eine deutsche Version wurden getestet, funktionieren und werden mit ausgeliefert.



Das Plug-In registriert eine Konfigurationsseite in der Kommandokonfiguration in der Sie das zu

verwendende AIML Set auswählen können.

simon wird die AIML Sets in den folgenden Ordnern suchen:

- GNU/LINUX: `` `kde4-config --prefix`/share/apps/ai/aimls/ ``
- Microsoft Windows: `[Installation Ordner (C:\Programme\simon 0.2\ per Default)]\share\apps\ai\aimls\`

Um die Persönlichkeit der Bots anzupassen, bearbeiten Sie bitte die Dateien bot.xml und vars.xml in den folgenden Ordnern:

- GNU/Linux: `` `kde4-config --prefix`/share/apps/ai/util/ ``
- Microsoft Windows: `[Installationsordner (C:\Programme\simon 0.2\ per default)]\share\apps\ai\util\`

Das Plug-In wird die mbrola Stimmen (http://tcts.fpms.ac.be/synthesis/mbrola.html) verwenden wenn diese installiert sind.

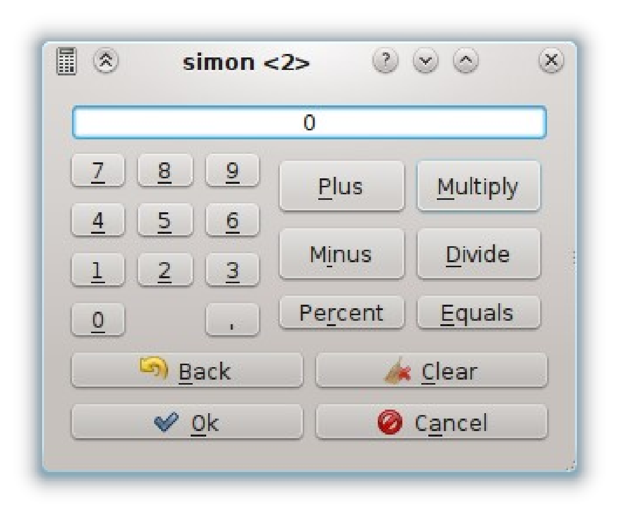
Weil keine Kommandos vom Typ "Diktation" angelegt werden können, scheint das Diktationsplug-in nicht im "Neues Kommando" Dialog auf.

Es wird empfohlen keinen Trigger für dieses Plug_in zu verwenden um die Gespräche "natürlicher" zu machen. Das AI plugin wird auf jedes Erkennungsergebnis in irgendeiner Form reagieren aber gleichzeitig - im Gegensatz zur Diktation - das Erkennungsergebnis an die anderen Plug-Ins weitergeben. Das ermöglicht es, mit speziellen AIML Sets Kommandos in eine natürliche Konversation einzubauen. Zum Beispiel könnte das KI Plug-In instruiert werden auf ein Erkennungsergebnis wie "Computer, Firefox" zu antworten mit "Firefox wird gestartet!"; da das Erkennungsergebnis ja aber zu den anderen Plug-Ins weitergegeben wird, wird auch ein mögliches Firefox-Kommando ausgeführt und der Browser tatsächlich gestartet.

Bitte beachten Sie, dass dieses Plug-In nur funktionieren wird, wenn festival installiert und konfiguriert ist und in Ihrem Systempfad liegt.

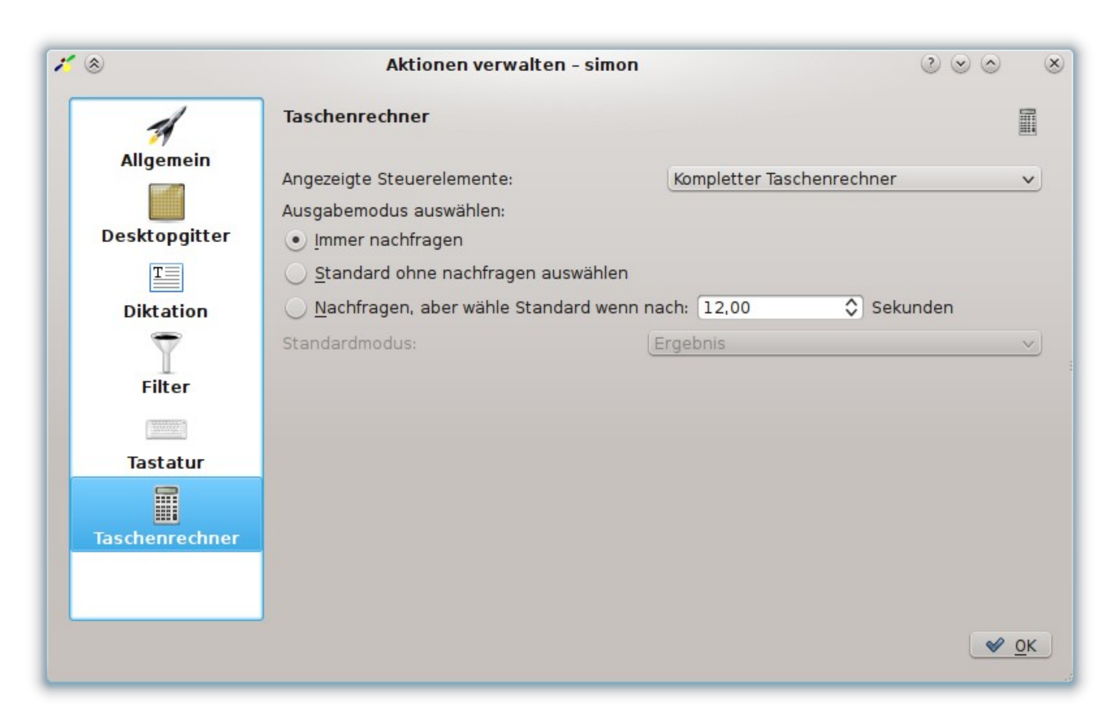
Taschenrechner

Das Taschenrechner Plug-In ist ein simpler, sprachgesteuerter Taschenrechner.

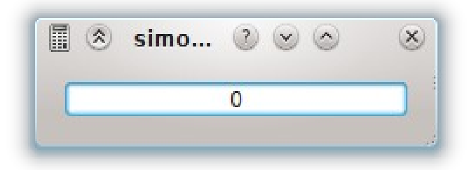


Der Taschenrechner erweitert das Nummer eingeben Plug-In mit weiteren Funktionen.

Wenn das Plug-In geladen wird, wird eine Konfigurationsseite zum Plug-Ins verwalten Dialog hinzugefügt.

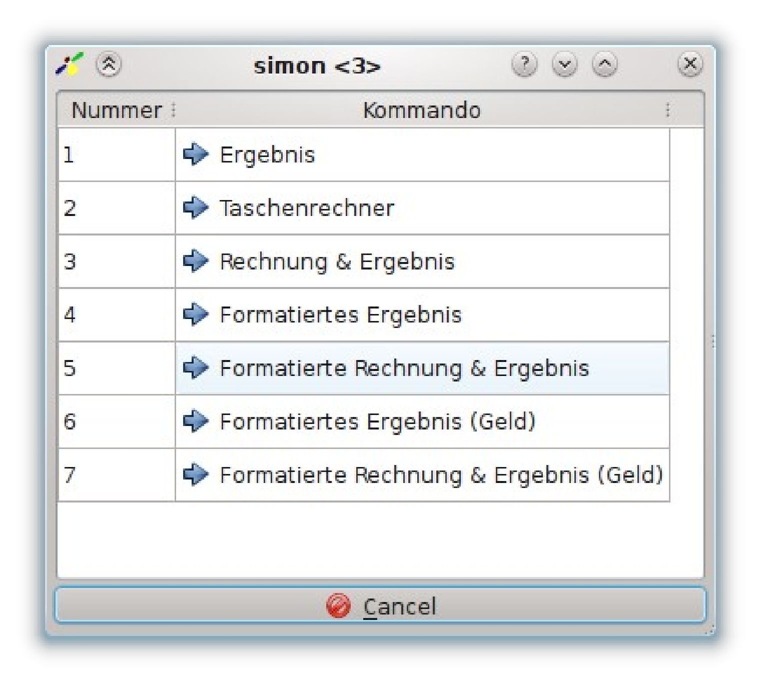


Dort können Sie auch den Kontrollmodus des Taschenrechner konfigurieren. Wird dieser auf etwas anderes als Voller Taschenrechner gesetzt werden Optionen ausgeblendet.



Die so versteckten Kontrollelemente können aber im Gegensatz zu solchen bei denen einfach alle assoziierten Sprachkommandos gelöscht wurden noch immer über die Sprache ausgewählt werden

Wenn der Ok ausgewählt wird, wird der Taschenrechner standardmäßig fragen was mit dem erzeugten Ergebnis gemacht werden soll. Sie können z.B. das Ergebnis der Rechnung, die Rechnung selbst oder beides ausgeben. Sie können dies jedes mal von der Liste auswählen oder in der Kommandokonfiguration fix setzen.



Filter

Mit dem Filter Plug-In können Sie Erkennungsergebnisse "abfangen" bevor Sie andere Plug-Ins erreichen. Mit diesem Plug-In können Sie beispielsweise die Erkennung über die Sprache pausieren.

Ein Filter hat zwei Zustäde:

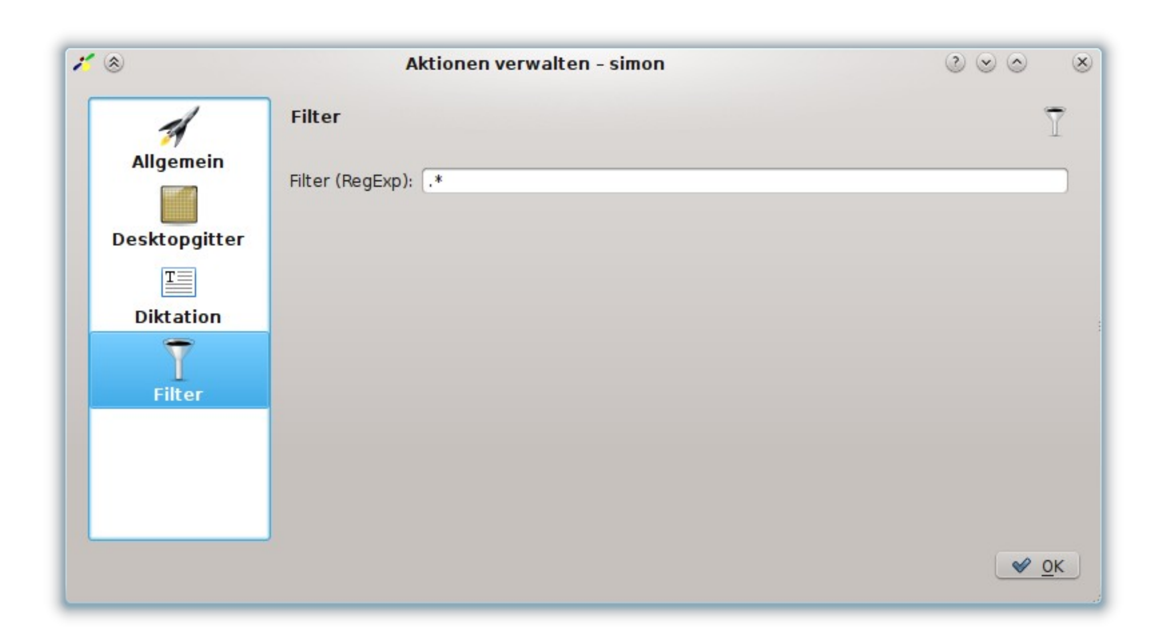
- Inaktiv

  Der Standardzustand. Alle Erkennungsergebnisse werden weitergereicht.

- Aktiv

  Wenn der Filter aktiviert ist, wird der Filter alle Ergebnisse "fressen" die zu einem konfigurierten String passen. Standardmäßig bedeutet dass, dass jedes Erkennungsergebnis vom Filter akzeptiert werden wird und somit nicht zu den folgenden Plug-Ins weitergereicht wird.

Das Filter Plug-In registriert eine Konfigurationsseite im Plug-Ins verwalten Dialog wo auch konfiguriert werden kann welche Ergebnisse herausgefiltert werden sollen.

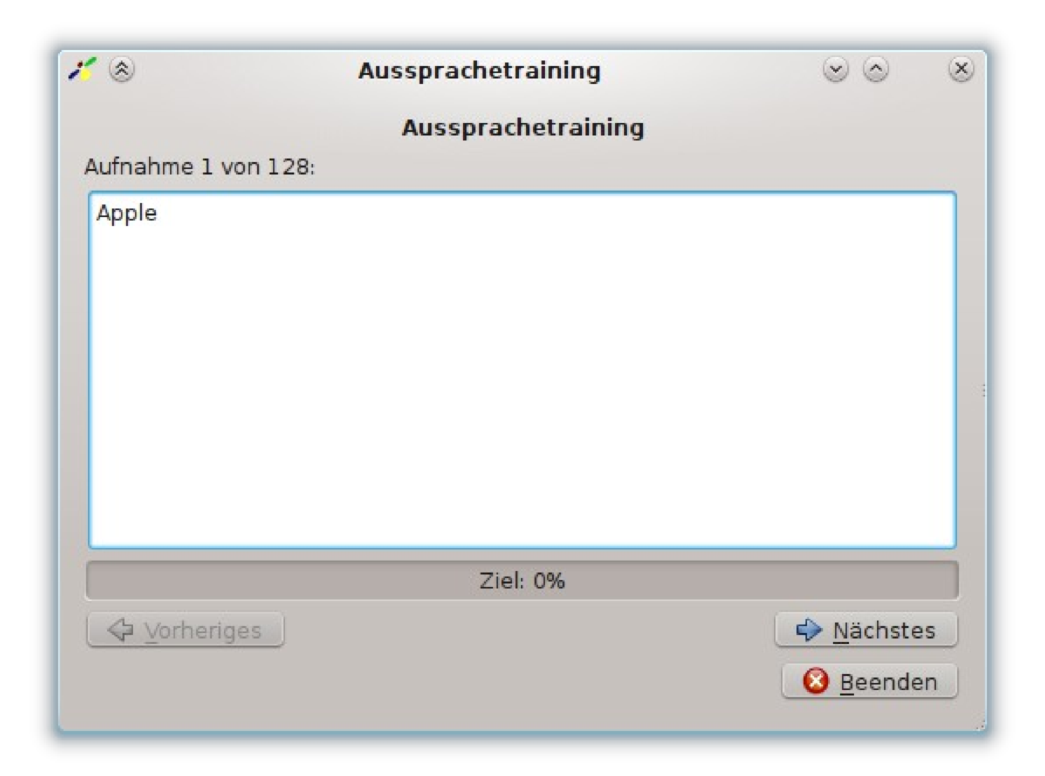


Der Filter ist ein regulärer Ausdruck der zur Laufzeit evaluiert wird sobald ein Erkennungsergebnis das Plug-In erreicht.

Das Plug-In registert Spracheingabekommandos zum aktivieren und deaktivieren des Filters.

Aussprachetraining

Das aussprachetraining, wenn kombiniert mit einem guten statischen Basismodell, kann ein hilfreiches Werkzeug sein um die Aussprache einer Fremdsprache zu trainieren.

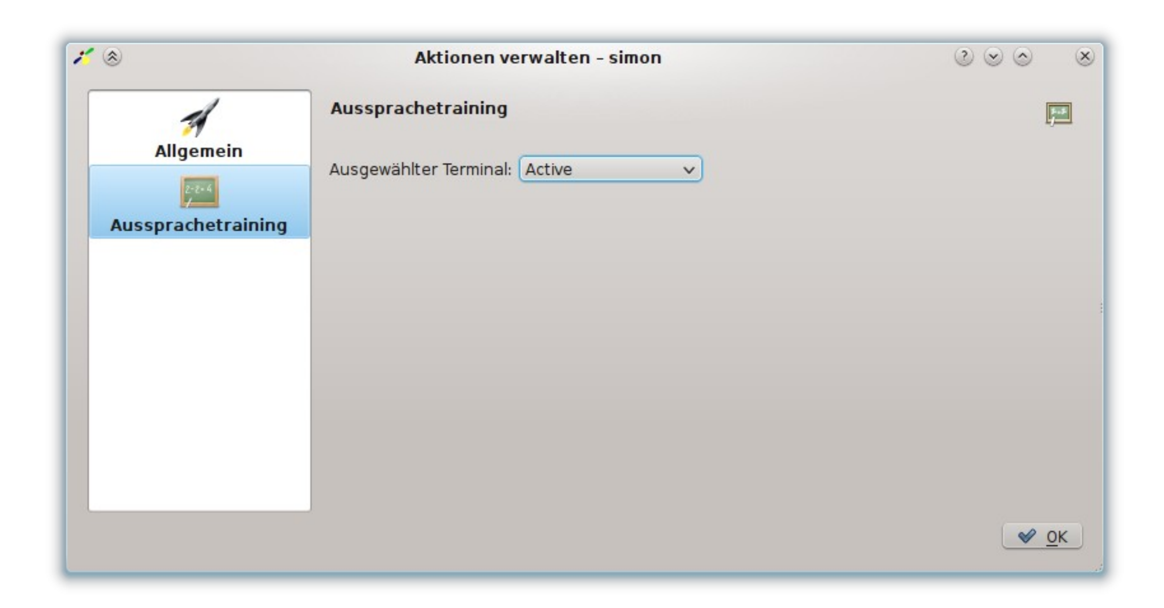


Das Plug-In wird Sie auffordern Wörter auszusprechen. Die Erkennung wird dann die Aufnahme analysieren und dabei je nach Aussprache hoffentlich das richtige Wort auswählen. Wird das richtige Wort erkannt wird anhand der Erkennungsgenauigkeit festgestellt wie nahe Ihre Aussprache an der der Sprecher des Basismodells liegt. Da Basismodelle in der Regel von Muttersprachlern aufgenommen wird, zeigt die angezeigte Erkennungsrate einen Qualitätsindex Ihrer Aussprache.

Je ähnlicher Sie zu den Muttersprachlern klingen, desto höher ist der angezeigte Wert.

Das Plug-In fügt einen Eintrag zum Kommandos Menü hinzu um das Training zu starten.

Das Training selbst besteht aus mehreren Seiten. Jede Seite beinhaltet ein Wort das vom aktiven Vokabular geholt wird. Die zu trainierenden Wörter werden per Terminal selektiert der in der Kommandokonfiguration eingestellt wird bevor das Training gestartet werden kann.



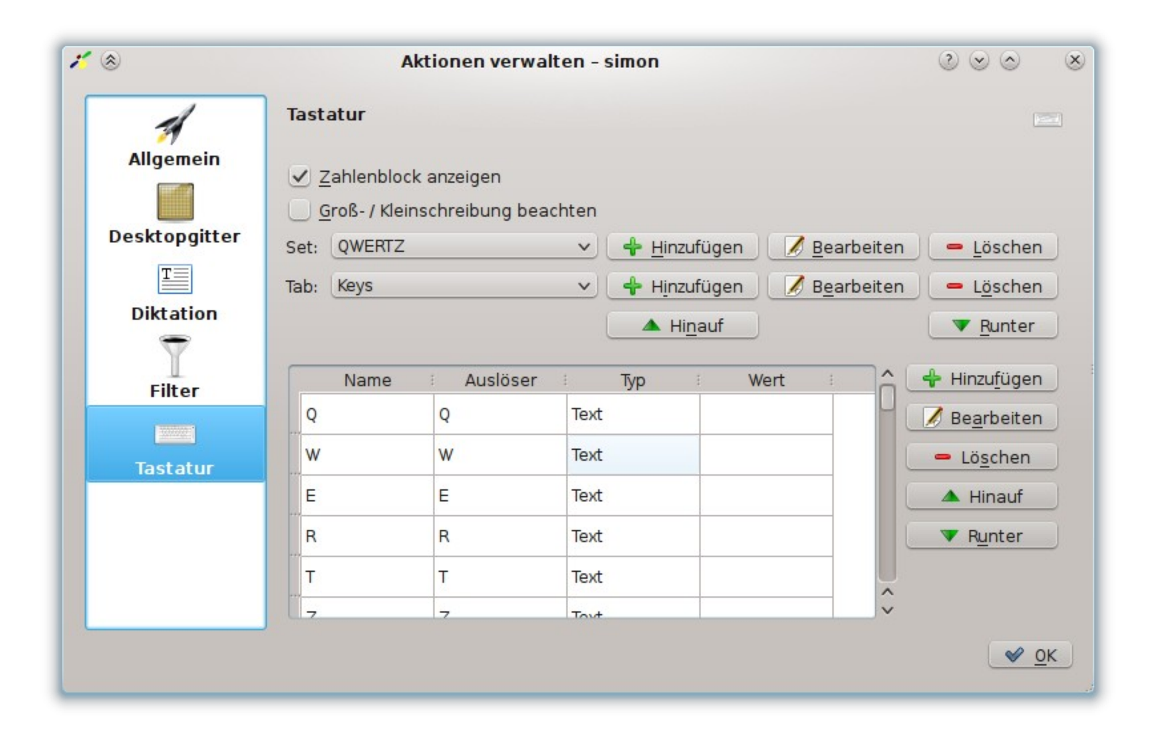
Tastatur

Das Tastatur Plug-In zeigt eine virtuelle, sprachgesteuerte Tastatur.



Die Tastatur besteht aus mehreren Registerkarten mit jeweils anderen Tasten. Die Gesamtheit von Registerkarten und Tasten sind in "Sets" gesammelt.

Sie können das zu verwendende Set in der Konfiguration auswählen oder neue anlegen.

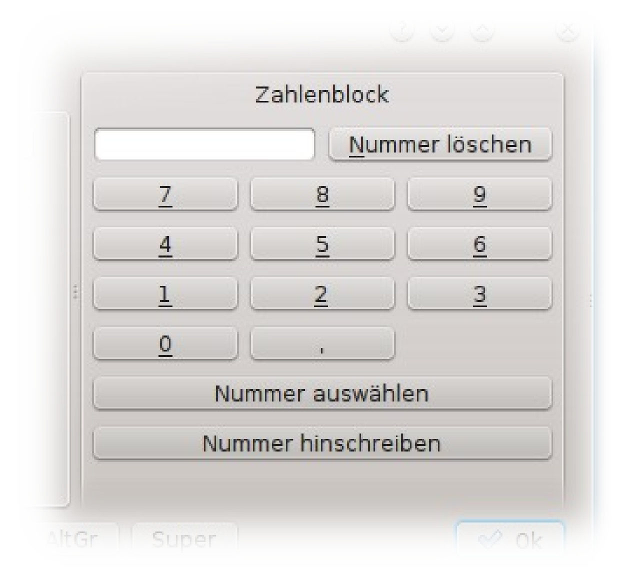


Tasten sind in der Regel mit genau einem Zeichen verbunden, können aber auch lägere Texte und sogar Tastenkürzel beinhalten. Deshalb können Tastatursets auch spezielle Tasten wie "Alle auswählen" oder "Passwort hinschreiben" besitzen.

Neben den Registerkarten die die Tasten aufnehmen kann die Tastatur einige spezielle Tasten wie Steuerung, Shift, etc. anzeigen. Diese Tasten sind als Spracheingabekommandos realisiert und werden unabhägig von der ausgewählten Registerkarte angezeigt.

Wie alle Spracheingabekommandos können die Knöpfe versteckt werden indem alle assoziierten Sprachkommandos entfernt werden.

Die Tastatur bietet außerdem einen Nummernblock der angezeigt werden kann indem die entsprechende Option inder Konfiguration gesetzt wird.



Neben den Nummerntasten existiert ein eigener Löschen-Knopf für das Nummereingabefeld (Nummer löschen) und zwei Knöpfe die bestimmen was mit der eingegebenen Nummer geschehen soll.

Wenn Nummer hinschreiben ausgewählt wird, wird die eingegebene Nummer mit simulierten Tastendrücken hingeschrieben. Nummer auswählen hingegen sucht einen Knopf der diese Nummer als Trigger konfiguriert hat und löst diesen aus. Auf diese Weise können Sie eine komplette Tastatur nur mit Nummern bedienen.



Die Tasten auf dem Nummernpad sind konfigurierbare Spracheingabekommandos

Konfiguration

simon ist auf extrem hohe Konfigurabilität ausgelegt.

Allgemein

Die allgemeine Konfiguration bietet einige basische Optionen.



Die Option simon beim Anmelden zu starten wird sowohl unter Microsoft Windows als auch unter KDE unter GNU/Linux funktionieren. Um andere Desktopumgebungen wie GNOME, XFCE, etc. zu unterstützen müssen Sie aber simon wahrscheinlich manuell im Sitzungs-Autostart platzieren (bitte sehen Sie die Handbücher Ihrer Desktopumgebung für mehr Informationen).

Wenn die Option zum minimiert Starten von simon gesetzt ist, wird das simon Hauptfenster beim simon Start nicht angezeigt sondern sofort in den Systembereich der Taskleiste minimiert.

Wenn Sie die option auswählen nicht zu warnen wenn Probleme mit Aufnahmen auftreten deaktiviert die Aufnahmen Qualitätskontrolle.

Aufnahmen

simon benutzt QtMultimedia um Ton aufzunehmen und wiederzugeben. QtMultimedia wird außerdem dazu verwendet einen Datenstrom vom Mikrofon an den Erkennungsserver zu senden.

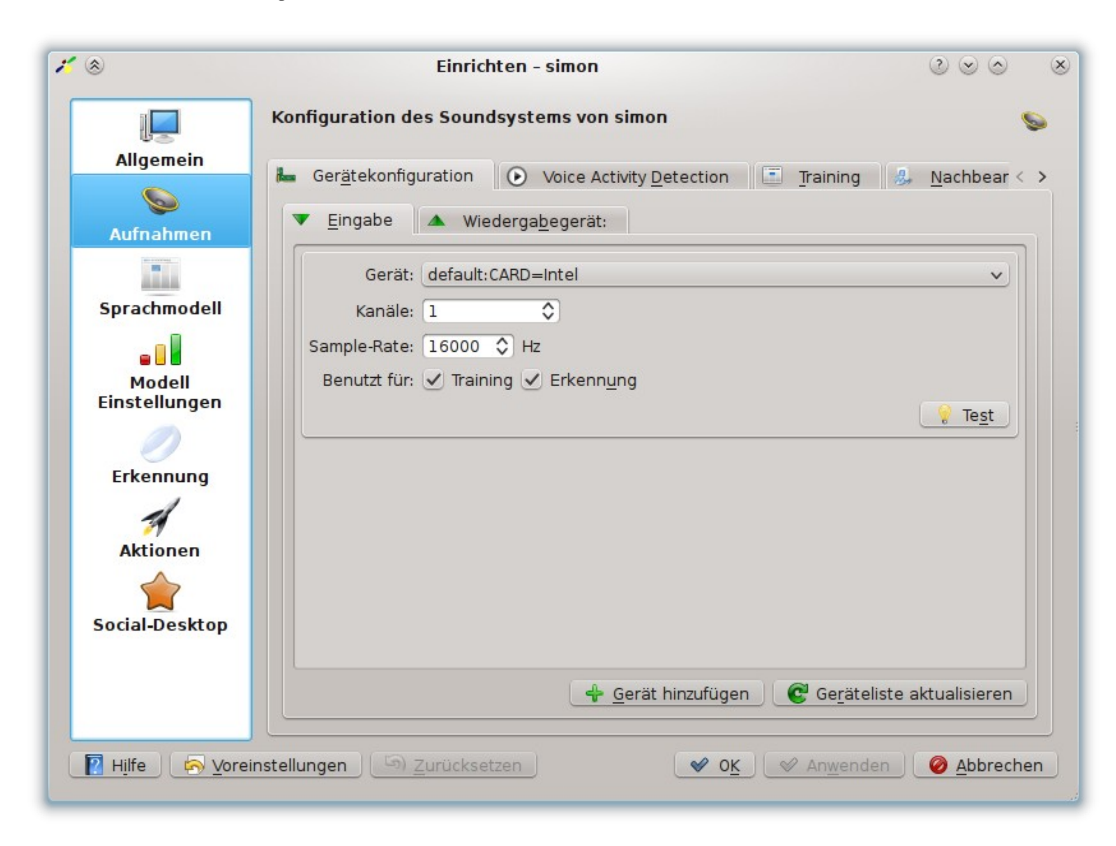
Gerätekonfiguration

Die Gerätekonfiguration erlaubt Ihnen das jeweilige Audiogerät auszuwählen das Sie für die Aufnahme und Wiedergabe in ganz simon verwenden möchten. Außerdem können sie auch angeben, in wievielen Kanälen und bei welcher Samplerate Sie aufnehmen wollen.

Der Defaultwert von einem Kanal und einer Samplerate von 16kHz sind normalerweise korrekt und müssen nicht geädert werden. Die Erkennung funktioniert nur mit mono Aufnahmen und funktioniert besser mit 16kHz Aufnahmen (8kHz wäre die Alternative).

Allerdings gibt es einige billige Audioadapter die diesen speziellen Aufnahmemodus möglicherweise nicht unterstützen. In diesem Fall kann simon mit Nachbearbeitungsketten und Software von Drittanbietern um diese Limitation herumarbeiten. Sehen Sie die Nachbearbeitungssektion für mehr Informationen.

Grundsätzlich gilt: Ändern Sie die Kanäle und die Samplerate nur wenn Sie sich sicher sind was Sie tun. Sonst wird die Erkennung *höchst wahrscheinlich* nicht funktionieren.



Verwenden Sie die Auswahllisten um die Geräte auszuwählen. Mit dem Knopf "Aktualisieren" können Sie simon veranlassen erneut nach Geräten zu suchen.

Sie können simon mit mehr als einem Audiogerät gleichzeitig verwenden. Verwenden Sie Gerät hinzufügen um ein Gerät zu Ihrer Konfiguration hinzuzufügen und Gerät entfernen um das Gerät wieder zu entfernen.

Das erste Gerät in Ihrer kann nicht entfernt werden.

Für jedes Gerät können Sie bestimmen wofür es verwendet werden soll: Training oder Erkennung (letzteres ist nur für Eingabegeräte verfügbar).

Wenn Sie mehr als ein Gerät für das Training verwenden wird simon mehrere Tondateien pro Trainingsaufnahme sammeln. Wird mehr als ein Gerät für die Erkennung verwendet werden mehrere Datenströme zum Server aufgebaut was in mehreren Erkennungsergebnissen (eines pro Mikrofon) resultiert.

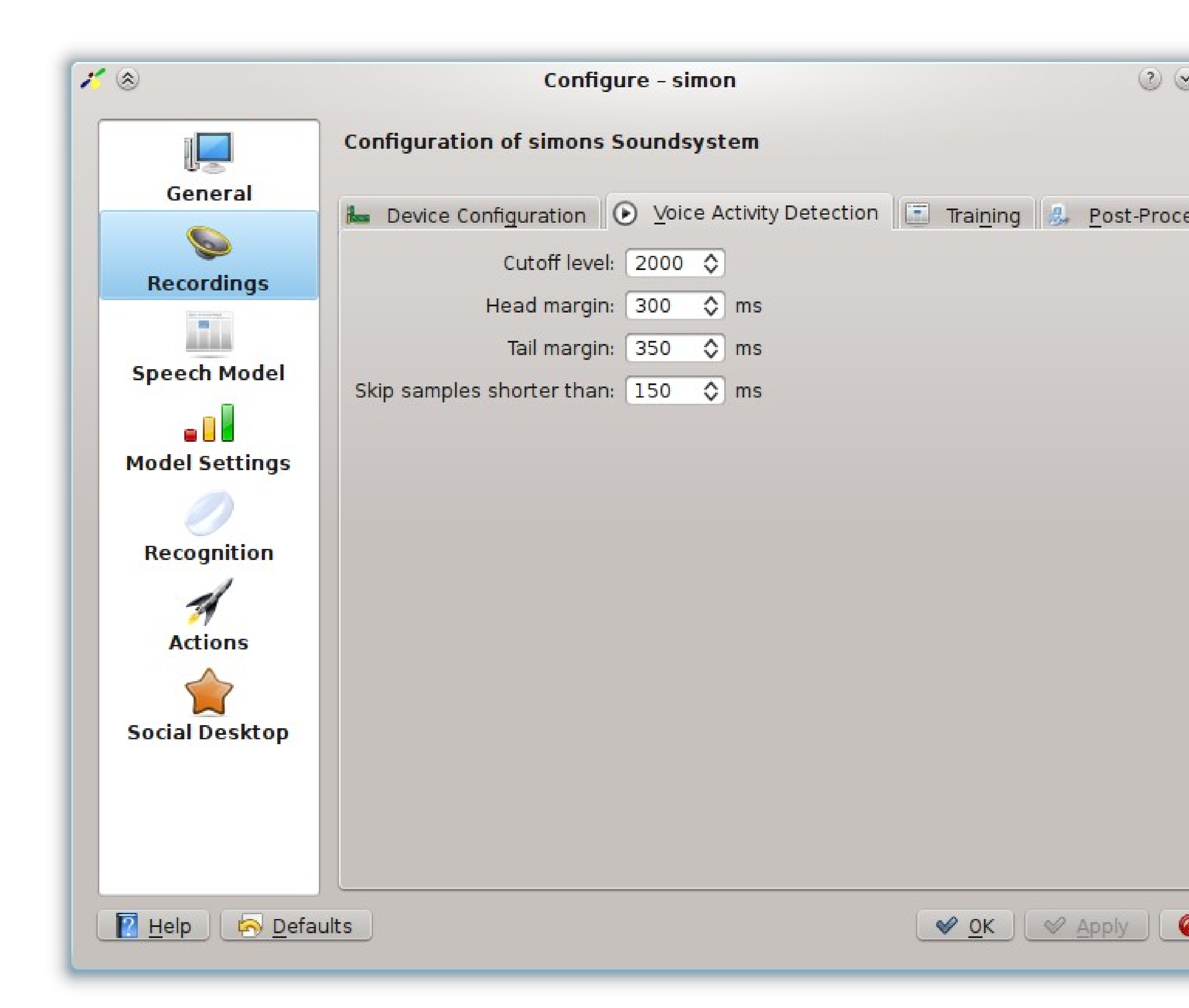
Wenn Sie mehr als ein Ausgabegerät verwenden wird ein Ton auf allen konfigurierten Geräten gleichzeitig wiedergegeben.

Wenn Sie unterschliedliche Samplerates für Ihre Eingabegeräte verwenden wird die Ausgabe nur auf diesen Ausgabegeräten stattfinden die ebenfalls die selbe Samplerate verwenden. Verwenden Sie beispielsweise ein Eingabegerät mit 16kHz und eines mit 48kHz wird die Wiedergabe von Aufnahmen die mit dem ersten Gerät aufgenommen wurden nur auf Ausgabegeräten stattfinden die auf 16kHz konfiguriert sind wohingegen Aufnahmen mit 48kHz nur auf Wiedergabengeräten abgespielt werden die selbst auf 48kHz eingestellt sind.

Sprachaktivitätserkennung

Die Erkennung selbst geschieht auf dem simond Server. Sehen Sie die Architektur Sektion für mehr Informationen).

Die Aufnahme-Daten vom Mikrofon werden nicht die ganze Zeit an den Server gesendet sondern werden vom simon Client segmentiert. Dies geschieht durch die sogenannte "Sprachaktivitätserkennung".



Hier können Sie die Segmentierung durch die folgenden Parameter beeinflussen:

- Cutoff level

  Alles unter dieser Lautstärke wird als "Stille" behandelt (Hintergrundgeräusch).

- Head margin

  Für die Dauer von Head Margin werden Daten gesammelt bevor die aktuelle Aufnahme tatsächlich als Sprache klassifiziert wird. Während dieser kompletten Zeit muss die Lautstärke über dem Cutoff Level bleiben.

- Tail margin

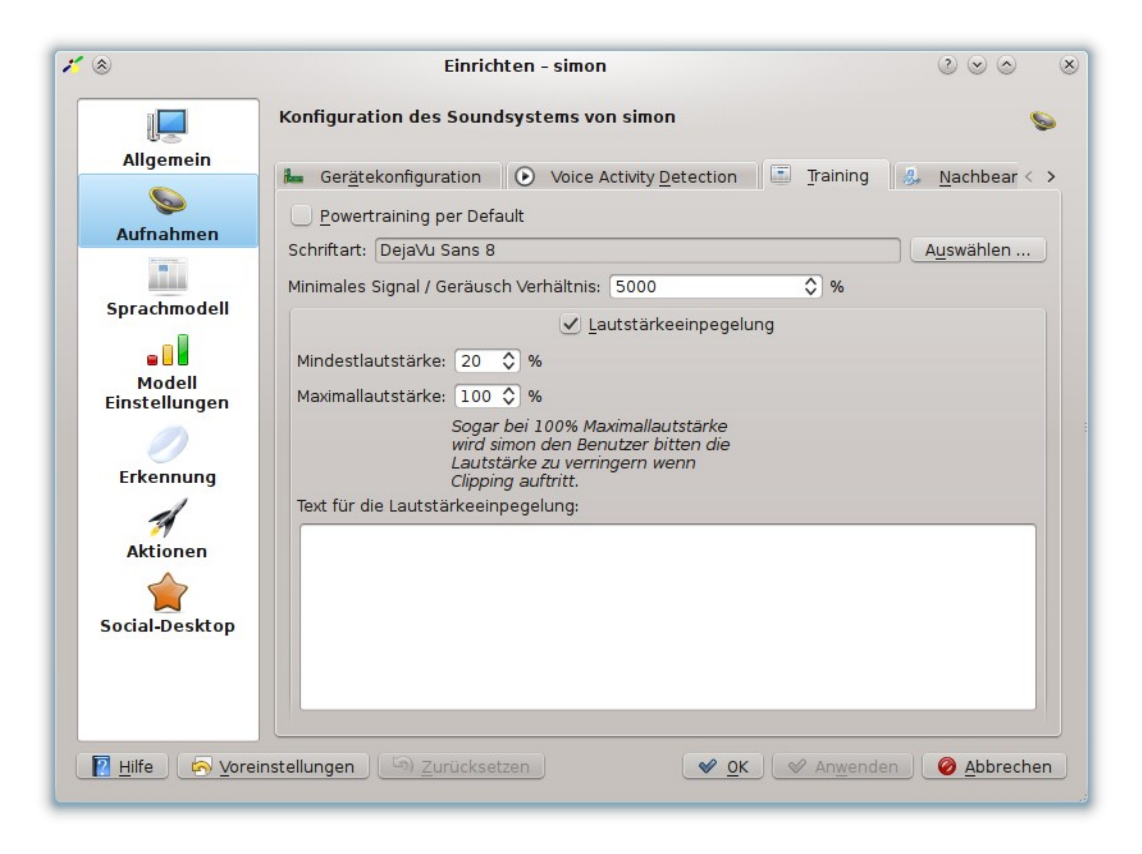
  Nachdem die Lautstärke unter das Cutoff Level fällt wird simon die Zeit von Tail Margin abwarten bis er die aktuelle Aufnahme abschließt.

- Aufnahmen überspringen die kürzer sind als

  Aufnahmen die kürzer als dieser Wert sind werden nicht an die Erkennung weitergegeben (husten,

etc.)

Trainingskonfiguration



Wenn die Option Powertraining per Default aktiviert ist wird simon beim Trainieren automatisch die Aufnahme starten bzw. stoppen wenn der dazugehörige Prompt angezeigt bzw. versteckt wird. Diese option setzt nur den Standardwert der Benutzer kann diesen bei jeder Trainingssitzung wieder überschreiben.

Der Text der aufgenommen werden soll um das Akustikmodell zu trainieren wird in einer konfigurierbaren Schriftart angezeigt. Die zu verwendende Schriftart und vor allem die Größe können hier geädert werden.

Diese Option wurde eingeführt, nachdem wir mit einigen Kunden Tests durchgeführt haben die unter infantiler Zerebralparese ("Spastische Lähmung") leiden. Während wir mit der Maus simon bedienten um sie durch das Training zu führen, mussten Sie im Rahmen des Trainings vorlesen was am Bildschirm geschrieben stand. Im Laufe des Tests fanden wir heraus, dass es unseren Probanden teilweise schwer fiel die kleine Schrift zu entziffern. Deshalb führten wir eine eigene Konfigurationsoption ein mit der eben jener Text der von den Probanden vorgelesen werden muss größer als der restliche Programmtext gemacht werden kann ohne den Rest des Programmes zu beeinflussen. Dies half außerdem um die Aufmerksamkeit direkt auf den vorzulesenden Text zu lenken.

Hier können Sie auch das benötigte Signal / Rauschverhältnis einstellen das erreicht werden muss damit eine Aufnahme von simon als korrekt angesehen wird. Sehen Sie die Aufnahme Qualitätskontrolle Sektion für mehr Details.

Auf dieser Konfigurationsseite können Sie auch die Paramter für die Lautstärkeeinpegelung festlegen.

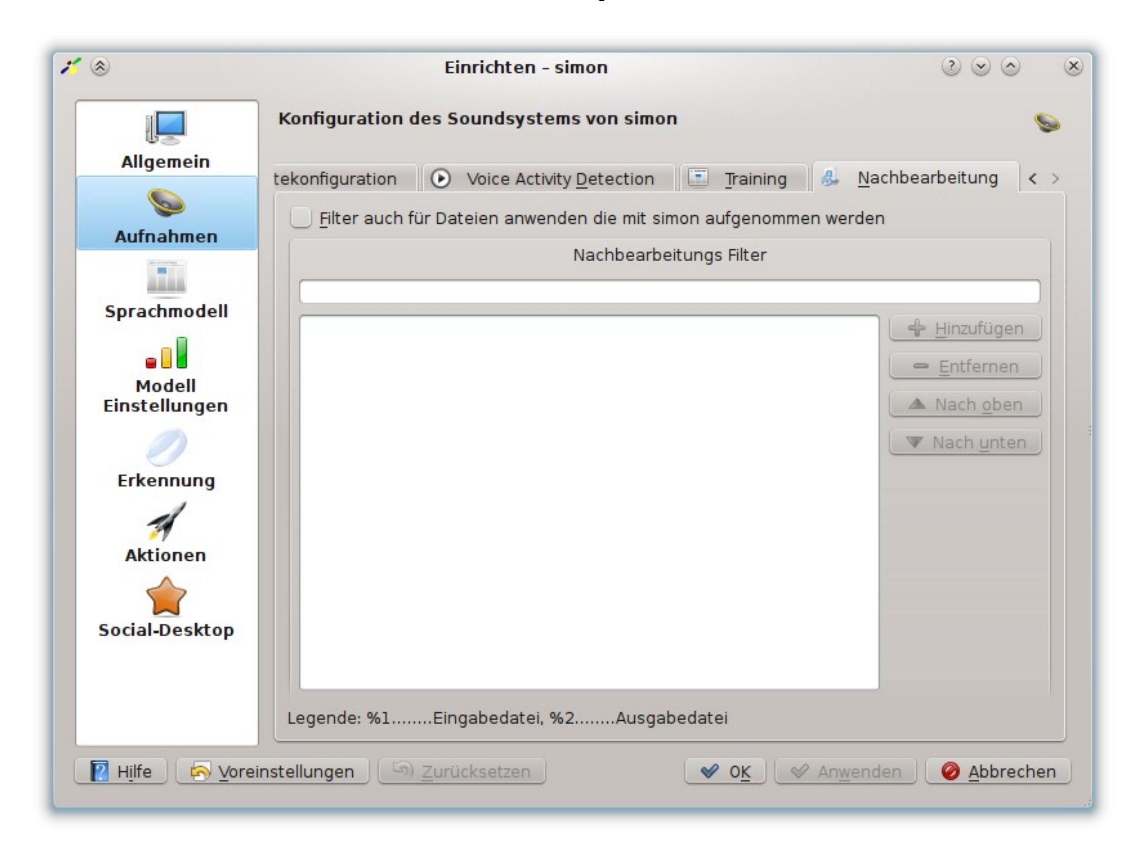
Sie kann sowohl für den Wort hinzufügen Dialog als auch für den Trainingsassistenten deaktiviert werden indem das Konfigurationsfeld deselektiert wird. Solange die aktuelle Aufnahmelautstärke unter der konfigurierten Minimumlautstärke bleibt wird der Anwender aufgefordert die Lautstärke zu erhöhen. Erreicht die Aufnahme die Maximallautstärke auch nur einmal wird simon Sie bitten die Lautstärke des Mikrofons zu verringern.

Übersteuern (erreichen der Maximallautstärke) wird immer eine "zu laut" Warnung erzeugen.

Der angezeigte Text kann konfiguriert werden indem Text im Eingabefeld eingetragen wird. Bleibt das Feld leer (Standard) wird ein Standardtext verwendet.

Nachbearbeitung

Alle aufgenommenen Sprachdaten (Training) und importierte Aufnahmen (mithilfe der Funktion "Trainingsdaten importieren") können mit einer Reihe von Nachbearbeitungskommandos bearbeitet werden. Nachbearbeitungsketten sind eine fortgeschrittene Funktionalität von simon die vom durchschnittlichen simon Anwender wohl nicht benötigt werden.



Die Nachbearbeitungskommandos sind wie eine Serie von Filtern durch die die Aufnahmen durch müssen. Mit diesen "Filtern" könnte man zum Beispiel Hintergrundrauschen entfernen oder die Aufnahmen normalisieren.

Nehmen wir als Beispiel ein Programm namens "audio_verarbeiten" das die Eingabe- und Ausgabedatei als Argumente nimmt (z.B.: "audio verarbeiten %1 %2". Die beiden Platzhalter %1 und %2 werden dann von simon bei jedem Aufruf mit den jeweiligen tatsächlichen Werten für den Eingabedateinamen

und den Ausgabedateinamen (respektive) ersetzt.

Die Option "Filter auch für Dateien anwenden die mit simon aufgenommen werden" definiert ob die konfigurierten Filter auch für mit simon erstellte Aufnahmen angewandt werden soll oder nur für Sprachdaten die mit der Funktion "Trainingsdaten importieren" importiert wurden.

Eine Beispielanwendung für Nachbearbeitungsketten wäre das Resampling von Aufnahmen auf Grund von Hardwareeinschräkungen. Nehmen wir eine Soundkarte an, die keine 16kHz mono Aufnahmen unterstützt sondern nur 44100Hz stereo ("CD") Aufnahmen, können wir das freie Kommandozeilentool SoX (http://sox.sourceforge.net) verwenden um die Aufnahmen automatisch auf ein entsprechendes Format zu resampeln.

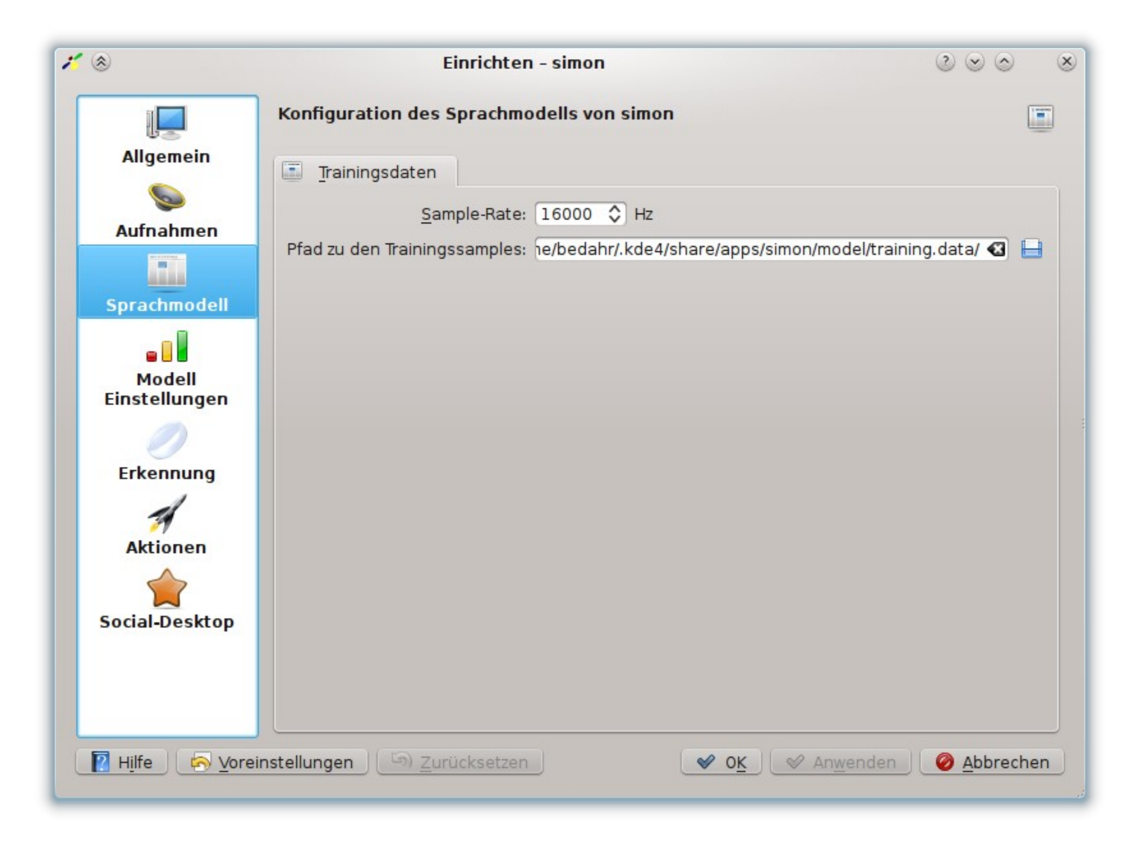
Dieses Beispiel würde das folgende Nachbearbeitungskommando verwenden:

- `sox -c 1 -r 16000 %1 %2`

Mit diesem Kommando können wir jetzt ohne Probleme mit 44100Hz und 2 Kanälen aufnehmen wenn die Option gesetzt ist, die Filter auch für Aufnahmen in simon anzuwenden weil simon dann die Aufgenommenen Dateien automatisch in brauchbare 16kHz mono Aufnahmen umwandeln wird wenn die Aufnahme abgeschlossen wird.

Sprachmodell

Hier können Sie die Parameter des Sprachmodells anpassen.



Die Samplerate die hier gesetzt wird ist die Ziel-Samplerate des Akustikmodells. Sie hat nichts damit zu tun in welcher Samplerate die Sprachdaten von simon aufgenommen werden. Es liegt in der

Verantwortung des Benutzers, dass die Sprachdaten schlussendlich für die Generierung des Modells auch in diesem Format vorliegen. Normalerweise wird das einfach erreicht, indem die Werte gleichgesetzt werden. Sollte dies nicht möglich sein, kann mit den Nachbearbeitungsketten nachgeholfen werden. Sehen Sie die Sektion Aufnahmen für mehr Details.

In den meisten Fällen werden entweder 16kHz oder 8kHz Modelle verwendet. 16kHz Modelle haben eine bessere Erkennungsrate als 8kHz Modelle und 8kHz Modelle sind weniger aufwädig (weniger Prozessor lastig) als 16kHz Modelle. Modelle mit mehr als 16kHz werden nicht empfohlen weil Sie sehr CPU-intensiv werden und in der Praxis kaum eine bessere Erkennungsrate erzielen.

Hier kann außerdem der Pfad zu den Trainingsdaten geädert werden. Sie müssen allerdings selbst sicherstellen, dass die bisher gesammelten Sprachaufnahmen auch in das neue Ziel verschoben werden. Alternativ würde simond die bereits synchronisierten Daten aus seiner Kopie wiederherstellen. Ein manuelles verschieben zum neuen Speicherort ist also nicht zwingend notwendig solange zuvor synchronisiert wird, aber es ist empfohlen.

Modellkonfiguration

Allgemein

Bitte sehen Sie die Basismodell Sektion.

Erweiterungen

Hier können Sie die Basis-URL konfigurieren die für den automatischen BOMP Import verwendet wird. Die Standardeinstellung zeigt auf die Kopie auf dem simon listens Server.

Erkennung

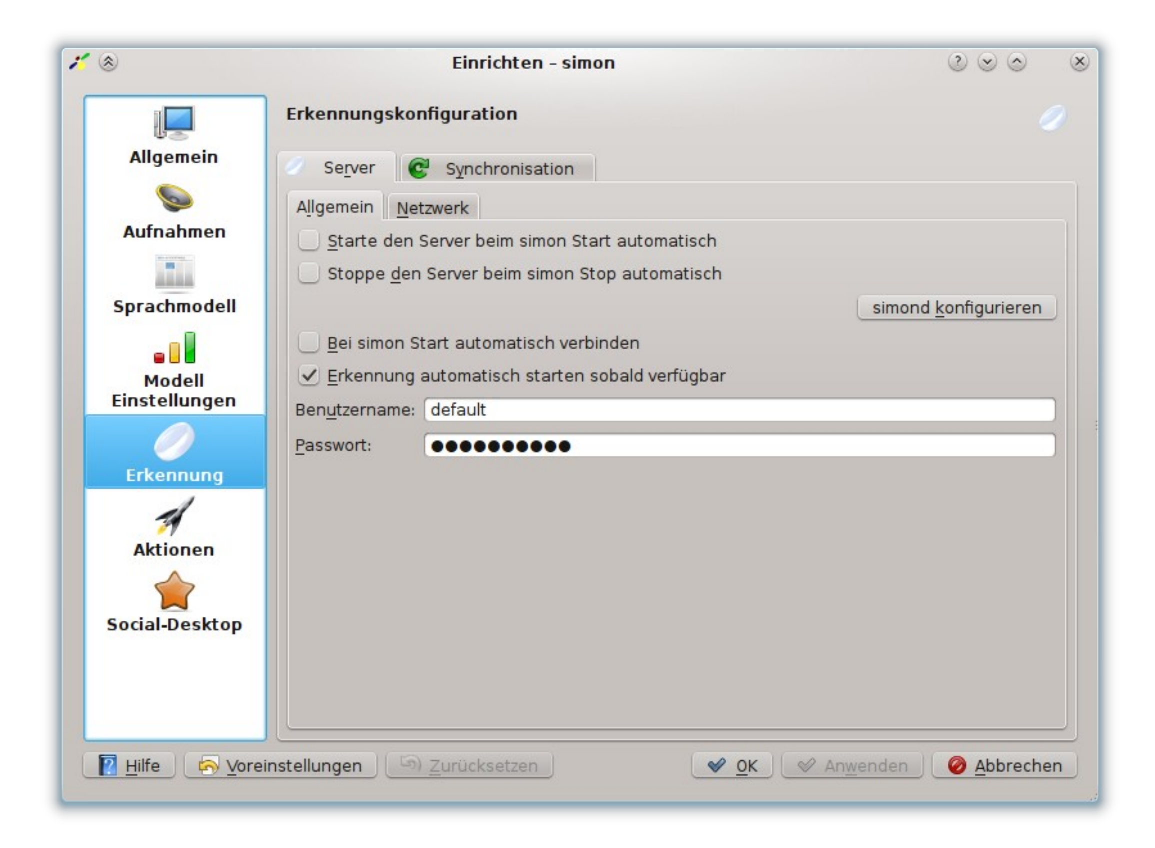
Hier können Sie die Erkennung und die Modellsynchronisation mit dem simond Server konfigurieren.

Server

Mit der Serverkonfiguration kann die Verbindung zum simond Server angepasst werden.

Allgemein

Die simon Hauptapplikation verbindet zum simond Server (siehe die Architektur Sektion für mehr Informationen).



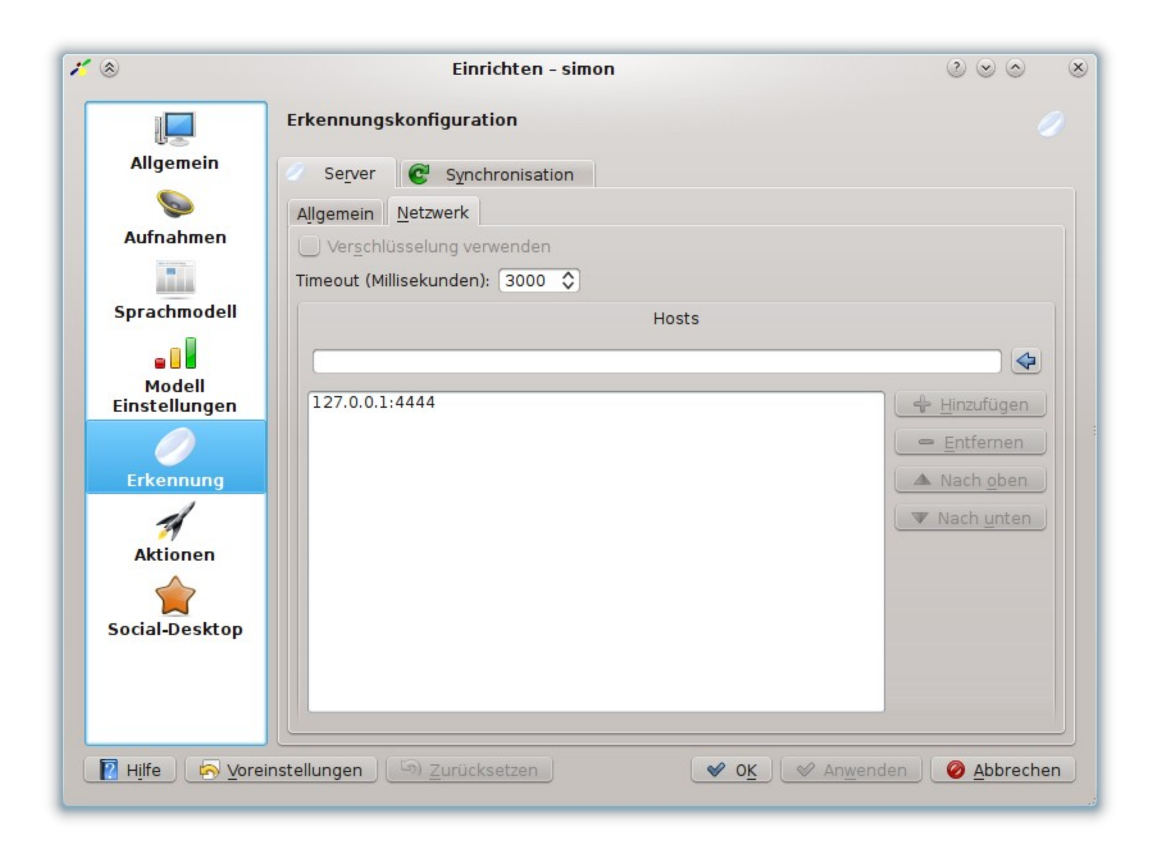
Um die einzelnen Anwender des Systems (ein simond Server kann natürlich mehrere simon Clients bedienen) zu identifizieren, führt simon Benutzer ein. Jeder Benutzer hat sein eigenes Sprachmodell. Die Benutzername / Passwort Kombination die hier angegeben werden kann, wird verwendet um sich damit bei simond anzumelden. Wenn der simond diesen Benutzernamen nicht kennt oder das Passwort nicht stimmt, wird die Verbindung abgebrochen werden. Sehen Sie das simond Handbuch (help:/simond) für mehr Information über das Einrichten der Benutzer auf dem simond Server.

Die Erkennung selbst - die ja vom Server übernommen wird - ist nicht immer verfügbar. Beispielsweise ist es nicht möglich die Erkennung zu starten bis der Benutzer sein Akustikmodell das erste mal erzeugt (was automatisch gemacht wird wenn bei der Synchronisation alle "Zutaten" vorliegen: Wortliste, Grammatik und Trainingsdaten). Per Default wird simon die Erkennung sofort starten sobald Sie verfügbar ist (Sprach- und Akustikmodell wurden konfiguriert). Dies kann hier geädert werden.

Mit der Option "Bei simon-Start automatisch verbinden" wird simon automatisch zum simond verbinden sobald die Applikation gestartet wurde.

Netzwerk

simon verbindet zum simond über TCP/IP.



Derzeit (simon 0.3) ist Protokollverschlüsselung nicht unterstützt.

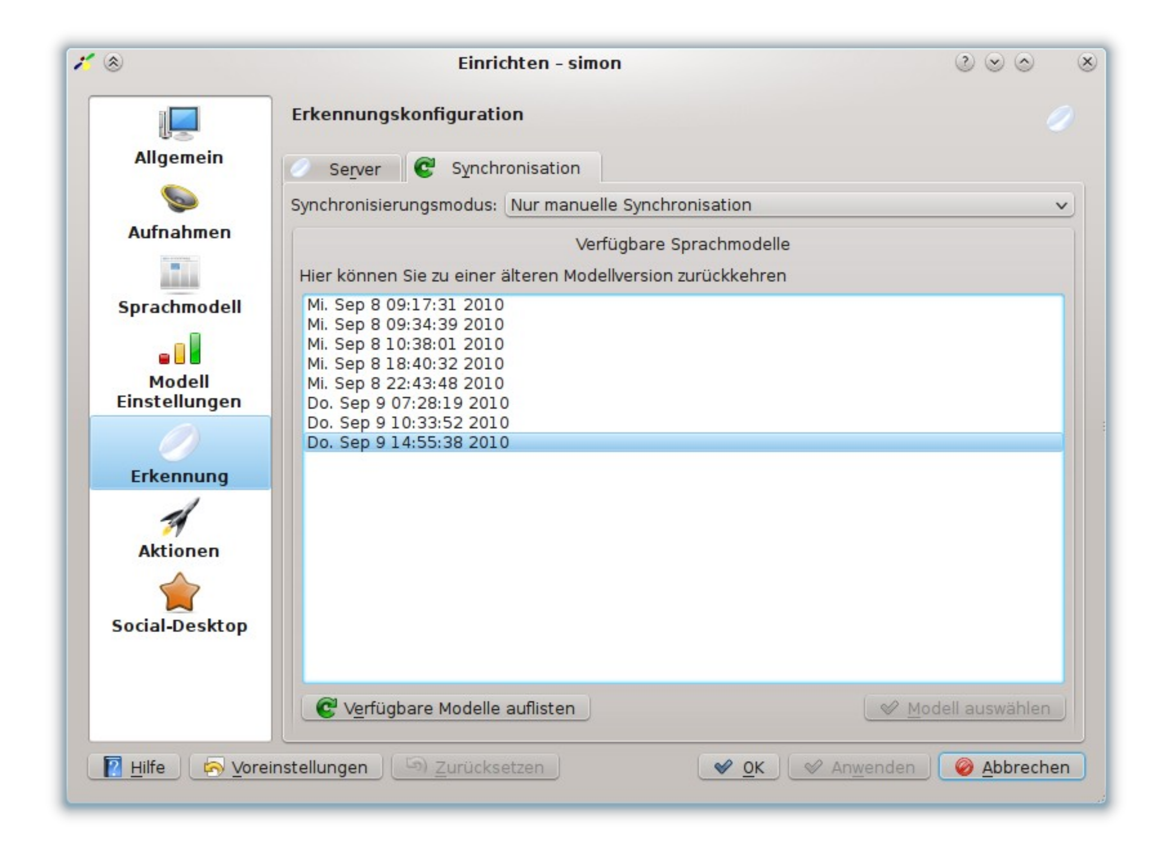
Der angegebene Timeout Wert gibt an wie lange simon auf eine Antwort der Gegenstelle wartet. Wenn Sie eine sehr, sehr langsame Netzwerkverbindung besitzen und/oder die "Beim simon-Start automatisch verbinden" Option verwenden und einen sehr langsamen Rechner verwenden erhöhen Sie diese Option wenn Sie Timeout Fehler beim Verbinden mit dem simond Server erhalten die durch erneutes Aufbauen der Verbindung behoben werden können.

simon unterstützt Konfigurationen bei denen mehr als ein simond angegeben wird. Das kann verwendet werden wenn Sie bspw. einen Laptop besitzen der unterschiedliche Server verwendet wenn er zu Hause oder bei der Arbeit verwendet wird. In diesem Fall können Sie beide Server hier angeben. Wenn simon verbindet, wird er die hier konfigurierten Server der Reihe nach durchprobieren bis er einen Server findet der auf seine Anfrage im gesetzten Timeout antwortet.

Um einen Server hinzuzufügen, geben Sie einfach seinen Hostnamen oder IP und den Port (mit ":" getrennt) ein oder verwenden Sie den Dialog der erscheint wenn Sie den blauen Pfeil neben dem Eingabefeld anklicken und wählen Sie "Hinzufügen".

Synchronisation und Modell Backup

Hier können Sie die Modellsynchronisation konfigurieren und wenn nötig ältere Versionen Ihres Sprachmodells wiederherstellen.



Erst nachdem ein Sprachmodell synchronisiert wurde werden alle Änderungen angewendet. Zusätzlich wird dabei auch ein neuer Wiederherstellungspunkt angelegt. Per Default wird simon nach jeder Änderung an den Modelldateien eine Synchronisation einleiten um die Änderungen zu übernehmen. Das nennen wir "automatische Synchronisation". Die automatische Synchronisation ist die empfohlene Einstellung.

Wenn Sei jedoch mehr Kontrolle über die Synchronisation wünschen, können Sie simon anweisen vor jeder Synchronisation nachzufragen oder die komplett manuelle Synchronisation auswählen bei der simon die Synchronisation nie automatisch einleiten wird - stattdessen müssen Sie den Menüpunkt "Synchronisation" verwenden (siehe Sektion das simon Hauptfenster) wenn Sie das Sprachmodell synchronisieren wollen.

Der simond Server wird eine Kopie der letzten fünf Versionen von den Modell-Eingabedateien (Wortliste, Grammatik, etc.) aufbehalten. Das fertige, erstellte Modell wird dabei nicht aufbehalten. Wenn Sie eine frühere Version einspielen wird aber das Modell anhand dieser Quelldateien automatisch neu generiert.

Wählen Sie dazu - wenn Sie mit dem Server verbunden sind - einen der verfügbaren Modelle aus und stellen Sie es mit einem Klick auf "Modell auswählen" wieder her.

Bitte beachten Sie, dass die Synchronisation generell nur komplette Quell-Modelle akzeptieren wird (eine Modellquelle ist komplett wenn Sie eine Wortliste, eine Grammatik und ein paar Trainingssamples besitzt) sodass nicht komplette Modelle nie auf dem Server gespeichert werden. Unvollstädige Quellmodelle obliegen damit auch nicht dem automatischen Backup.

Aktionen

Mit der Aktionskonfiguration kann die Reaktion auf Erkennungsergebnisse konfiguriert werden.

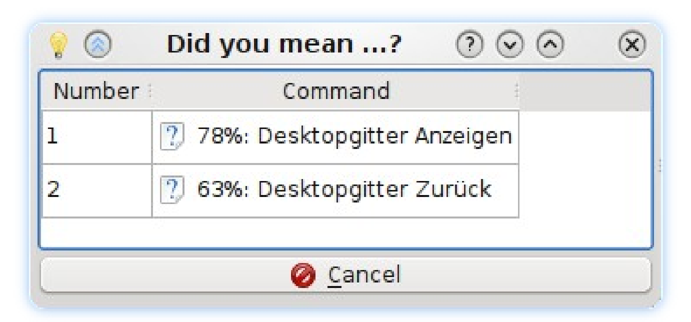
Erkennung

Neben dem wahrscheinlichsten Ergebnis kalkuliert simon die 10 besten Erkennungsergebnisse.

Jedes Ergebnis wird mit einer Punkteanzahl zwischen 0 und 1 versehen (wobei 1 einer Sicherheit von 100% entspricht).

Mit Hilfe der Minimale Sicherheit können Sie eine minimale Sicherheit für Erkennungsergebnisse unter dieser Ergebnisse ignoriert werden.

Wenn mehr als ein Erkennungsergebnis diesen Wert überschreitet wird simon ein Popup anzeigen das die wahrscheinlichsten Ergebnisse anzeigt aus denen dann ausgewählt werden kann.

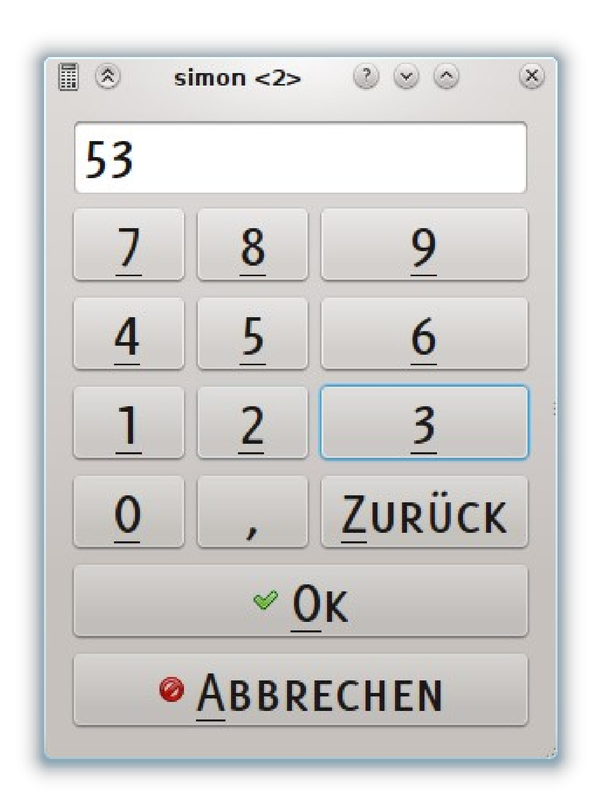


Dieses Pup kann über die Option Auswahlliste anzeigen für ungenaue Ergebnisse kontrolliert werden.

Plug-In Basisschriftart

Einige Plug-Ins von simon haben eine grafische Oberfläche.

Die Schriftarten dieser Oberflächen können hier zentral konfiguriert werden - unabhägig von den Systemschriftarten Ihres Betriebssystems.



Listen

Hier finden Sie die globale Listenkonfiguration. Diese fungiert als Vorlage für neue Szenarios aber wird auch direkt verwendet für das Auswahlfenster für mehrdeutige Erkennungsergebnisse.

Sozialer Desktop

Szenarios können hochgeladen und heruntergeladen werden mt simon.

Dafür verwenden wir KDEs soziale Desktop Fähigkeiten und unsere eigene Kategorie für simon Szenarios auf kde-files.org (http://kde-files.org/index.php?xcontentmode=692).

Wenn Sie bereits einen Account auf opendesktop.org (http://opendesktop.org) besitzen, können Sie die Anmeldedaten hier angeben. Ansonsten können Sie direkt im Konfigurationsmodul auch einen neuen Account anlegen.

Die Registrierung ist natürlich kostenlos.

Aufnahmeparameter manuell anpassen

simon ist auf Endbenutzer ausgelegt. Das Benutzerinterface ist darauf ausgelegt, es sogar Endbenutzern ohne Hintergrund in Spracherkennung zu ermöglichen eigene Sprach- und Akustikmodelle zu designen indem viele Default-Werte bereits beim ersten Öffnen der Applikation festgelegt wurden.

In speziellen Fällen (zum Beispiel schwere Sprachstörungen) kann es notwendig werden, einige tiefgreifende Parameter anzupassen. Deshalb respektiert simon auch die gesetzten Werte in den Konfigurationsdateien für die Erkennung. Diese Konfigurationsdateien können natürlich frei angepasst

werden als würde man direkt mit Julius arbeiten.

Es gibt zwei Teile der Julius Konfiguration die angepasst werden können:

- adin.jconf

  Das ist die Konfiguration die den Soundstream vom simon Client zum simond server steuert. Die Datei wird direkt vom adinstreamer gelesen.

  simon liefert eine Defaultdatei aus ohne spezielle Parameter. Sie können die Systemweite Konfiguration anpassen die alle Benutzer betrifft wenn Sie mehr als einen simon Benutzer auf Ihrem physischen System haben. Wenn Sie die Konfiguration nur für einen speziellen Benutzer anpassen wollen, kopieren Sie die Konfiguration zum Benutzerpfad und passen Sie die Kopie an.

- julius.jconf

  Das ist die Konfiguration vom simond Server die direkt die Erkennung beeinflusst. Die Datei wird direkt von libjulius und libsent gelesen.

  simond wird auch mit einer Default `julius.jconf` ausgeliefert. Immer wenn ein neuer Benutzer zur simond Datenbank hinzugefügt wird, wird simond automatisch eine Kopie der Systemweiten `julius.jconf` anlegen und sie in den Ordner des neuen Benutzers kopieren. Sie können die einzelnen Dateien der jeweiligen simond Benutzer natürlich frei anpassen. Die Änderungen am systemweiten "Template" beeinflussen dadurch aber keine bereits angelegten Benutzer.

Der Pfad zu den Julius Konfigurationsdateien wird von der verwendeten Plattform abhägen:

Table 4-5. Julius Konfigurationsdateien

Datei	Microsoft Windows	GNU/Linux
adin.jconf (System)	(Installationspfad)\share\apps\simon\adin.jconf	`kde4-config --prefix`/share/apps/simon/adin.jconf
adin.jconf (Benutzer)	%appdata%\.kde\share\apps\simon\adin.jconf	~/.kde/share/apps/simon/adin.jconf
julius.jconf (Template)	(Installationspfad)\share\apps\simond\default.jconf	`kde4-config --prefix`/share/apps/simond/default.jconf
julius.jconf (Benutzer)	%appdata%\.kde\share\apps\simond\models\(Benutzer)\active\julius.jconf	~/.kde/share/apps/simond/models/(Benutzer)/active/julius.jconf

Chapter 5. Fragen und Antworten

Um diese Sektion immer aktuell zu halten, wird dieser Teil des Handbuches in unserem online Wiki (http://www.cyber-byte.at/wiki/index.php/English:_Troubleshooting) und somit für alle editierbar verwaltet.

Chapter 6. Danksagungen und Lizenz

simon

Appendix A. Installation

Bitte sehen Sie unser online Wiki (http://www.cyber-byte.at/wiki/index.php/English:_Setup) for eine detaillierte Installationsanleitung.